PROPERTIES OF GENERALIZED DEFINITIONS OF LIMIT*

RALPH P. AGNEW

1. **Introduction.** The theory of summability has been the subject of several excellent expository addresses† presented to this Society. These addresses have dealt largely with properties of matrix transformations

$$A: \qquad y_s = \sum_{t=0}^{\infty} A_{s,t}x_t,$$

which associate with certain sequences $x_0, x_1, x_2, \cdots$ of complex numbers the sequences $y_0, y_1, \cdots$ determined by use of a given matrix $A_{s,t}$ of complex constants.

It is my object to discuss, and to compare with the matrix transformations A, the kernel transformations

$$K: \qquad y(s) = \int_0^{\infty} K(s, t)x(t)dt,$$

which associate with certain complex-valued functions $x(t)$ defined over $0<t<\infty$ the functions $y(t)$ determined by a given kernel $K(s, t)$ belonging to a certain class of complex-valued functions which we specify in §3. Transformations of this form were first studied by Silverman.‡ More recent contributions§ have been made by Knopp, Hill, Raff, and Day.

The point of view of the present study of kernel transformations is quite different from that of earlier ones. The earlier studies have started with either the Riemann or Lebesgue integral and the class X

* An address delivered before the New York meeting of the Society on February 25, 1939, by invitation of the Program Committee.

† W. B. Ford, this Bulletin, vol. 25 (1918–1919), pp. 1–15; R. D. Carmichael, ibid., vol. 25 (1918–1919), pp. 97–131; C. N. Moore, ibid., vol. 25 (1918–1919), pp. 258–276; W. A. Hurwitz, ibid., vol. 28 (1922), pp. 17–36; and C. N. Moore, ibid., vol. 37 (1931), pp. 240–250.

‡ L. L. Silverman, *On the notion of summability for the limit of a function of a continuous variable*, Transactions of this Society, vol. 17 (1916), pp. 284–294.

§ K. Knopp, *Zur Theorie der Limitierungsverfahren*, Mathematische Zeitschrift, vol. 31 (1929–1930), pp. 97–127; pp. 276–305. J. D. Hill, *A theorem in the theory of summability*, this Bulletin, vol. 42 (1936), pp. 225–228. H. Raff, *Lineare Transformationen beschrankter integrierbarer Funktionen*, Mathematische Zeitschrift, vol. 41 (1936), pp. 605–629; *Über lineare Integraltransformationen*, Monatshefte für Mathematik und Physik, vol. 45 (1937), pp. 379–393. M. M. Day, *Regularity of function-to-function transformations*, this Bulletin, abstract 44-9-332.

of all functions $x(t)$ which are bounded and integrable over each finite interval; and the fundamental problem solved is that of characterizing the functions $K(s, t)$ defining transformations *regular over the assigned set* X, that is, transformations such that each $x \in X$ for which $\lim_{t\to\infty} x(t)$ exists has a transform $y(s)$ for which $\lim_{s\to\infty} y(s) = \lim_{t\to\infty} x(t)$. The papers of Knopp are apparent exceptions; but later writers have implied that Knopp should have introduced the set X to make his work precise. The present study starts with a definition of integral (see §2) and a kernel $K(s, t)$ belonging to a certain class of functions (see §3); and the fundamental problem which presents itself is that of setting up criteria to determine what properties the transformation thereby determined has or fails to have.

This address is entirely self-contained in the sense that no knowledge of either the now extensive known theory of matrix transformations, or the more modest known theory of kernel transformations, is assumed, and that all proofs are given in terms of fundamental notions of analysis. That this is so is not purely a recognition of the fact that this address should be so constructed to meet the needs of optimistic individuals who, without previous experience with the theory of summability, might hope to gain from this address some knowledge of the theory. The author has felt for years that those who work in the theory of summability (and in particular the author himself) should have in print a self-contained foundation for further work in the theory of kernel transformations.

Some examples and remarks indicate the manner in which we begin to develop de novo the theory of kernel transformations. Two of the simplest and most useful transformations of the forms A and K respectively are

$$y_s = \sum_{t=0}^{s} \frac{1}{s+1} x_t, \qquad s = 0, 1, 2, \cdots, \tag{1.01}$$

in which $A_{st} = 1/(s+1)$ or 0 according as $0 \leqq t \leqq s$ or $t > s$; and

$$y(s) = \int_0^s \frac{1}{s} x(t)dt, \qquad s > 0, \tag{1.02}$$

in which $K(s, t)$, defined for $s > 0$, is $1/s$ or 0 according as $0 \leqq t \leqq s$ or $t > s$. We shall emphasize later the point that the transformation (1.02) does not become meaningful until the definition of integral used there has been specified. Let us use for the moment either the Lebesgue integral or the "improper" Cauchy-Riemann integral. If we let the function $y(s)$ given by (1.02) be denoted by $y_1(s)$, and let the

(1.02) transform of $y_{r-1}(s)$ be denoted by $y_r(s)$, it can easily be shown by induction that, for each $r=1, 2, \cdots$,

$$y_r(s) = \frac{1}{\Gamma(r)s}\int_0^s \left(\log\frac{s}{t}\right)^{r-1} x(t)dt. \tag{1.03}$$

As a matter of fact (1.03) defines, for each complex r having a positive real part, a transformation (the Hölder transformation of order r) with kernel

$$\begin{aligned} K(s, t) &= [\log (s/t)]^{r-1}/\Gamma(r)s, & 0 < t < s, \\ &= 0, & t \geqq s. \end{aligned} \tag{1.04}$$

The orthodox (ϵ, δ) definition of limit is, from a sufficiently abstract point of view, one scheme for associating with each sequence x_s, or function $s(x)$ belonging to a certain class, a number L called its limit. The transformations A and K furnish *generalized definitions of limit* or *methods of summability* when one defines $\lim_{s\to\infty} y_s$ and $\lim_{s\to\infty} y(s)$ to be generalized limits of a sequence x_s and a function $x(s)$, respectively, when the limits exist. We conform to accepted terminology in calling x_s *summable* A to L in case $y_0, y_1, \cdots$ exist and $\lim_{s\to\infty} y_s = L$, and $x(s)$ *summable** K to L in case $y(s)$ exists for $s>0$ and $\lim_{s\to\infty} y(s) = L$. Thus each method of summability furnishes, as does the (ϵ, δ) definition of limit, a scheme of associating with certain sequences or functions numbers which may be called their "limits."

A transformation A (or matrix A) is called regular if each convergent sequence x_t is summable A to the value to which it converges. For example, it is well known and is a good exercise for undergraduates to show that (1.01) is regular. Formulation of a useful definition of regularity of K is not quite so simple, and is postponed to §5.

Let $A_{s,t}$ be a matrix and let x_t and y_t be sequences so related that

$$y_s = \sum_{t=0}^{\infty} A_{s,t}x_t, \qquad s = 0, 1, 2, \cdots. \tag{1.05}$$

If step functions $x(t)$ and $y(s)$ and a step kernel $K(s, t)$ are defined by the equations

$$x(t) = x_{[t]}, \quad y(s) = y_{[s]}, \quad K(s, t) = A_{[s],[t]}, \tag{1.06}$$

where $[r]$ denotes the greatest integer less than or equal to r, then

* It is of course possible to modify these definitions, calling x_s or $x(s)$ summable A or K to L in case y_s or $y(s)$ exist for all sufficiently great s and $\lim_{s\to\infty} y_s = L$ or $\lim_{s\to\infty} y(s) = L$. These modified definitions, which turn out to be not significantly different from ours, are discussed at the end of §9.

under any one of several definitions of integral (1.05) can be written

$$y(s) = \int_0^\infty K(s, t)x(t)dt, \qquad s > 0. \tag{1.07}$$

This means simply that any matrix transformation A can be represented as a kernel transformation K in which the domain and range are limited to step functions constant over each interval $n \leqq t < n+1$, $n = 0, 1, \cdots$. On account of this fact (which will be discussed in more detail in §10), an advance in the theory of matrix transformations suggests at least the possibility of making a corresponding advance in the theory of kernel transformations.

The theory of kernel transformations K has lagged far behind the theory of matrix transformations A. This is unquestionably due in part to the fact that the theory of K is really more difficult than that of A. The author feels that this is also due in part to the fact that in spite of much work on linear transformations in general and kernel transformations in particular, there never has been an adequate foundation laid for development of the theory of K analogous to recent developments in the theory of A.

When we compare a kernel transformation K with a matrix transformation A, a fundamental difference between the two appears immediately. On the one hand,

$$\sum_{t=0}^{N} A_{s,t}x_t = A_{s,0}x_0 + A_{s,1}x_1 + \cdots + A_{s,N}x_N \tag{1.08}$$

always exists and has a unique meaning for all mathematicians when a matrix A, a sequence $x_1, x_2, \cdots$ of complex numbers, and an integer $N \geqq 0$ are given. On the other hand, specification of a kernel $K(s, t)$, a function $x(t)$, and a number $h > 0$ is not (in these days when a multiplicity of different definitions of integral are used in analysis) sufficient to determine whether

$$\int_0^h K(s, t)x(t)dt \tag{1.09}$$

exists. Not only the fact of existence of (1.09) but also the value of (1.09) may depend upon the definition of integral used.

It is neither economical nor satisfying to develop the theory of kernel transformations first for Riemann integrals, then for Lebesgue integrals, then for one or more other integrals, and then perhaps finally for general integrals of Banach type under which each bounded function is integrable over each bounded set. Moreover an attempt to

study kernel transformations without prescribing the type or the properties of the integrals involved is utterly futile and meaningless.

It turns out that a considerable part of a theory of K analogous to known theory of A can be developed for any definition of integral having the eight properties which we give in §2. In §3, we define *kernel* and in §4 we discuss the class $\mathfrak{K}$ of functions $x(t)$ such that the integral

$$\int_0^h K(s, t)x(t)dt$$

exists for each pair of positive numbers h and s. In §5, regularity is defined and discussed. In §§6–8, we give several theorems which supply the tedious parts of proof of necessity for theorems on regularity, and so on, in §9. The theorems of §§6–8 are made sufficiently general to furnish proof of necessity for many other theorems in the theory of summability which we shall be unable to give in a paper of temperate length. In §10, transformations whose kernels are step kernels are related to sequence-to-function and matrix transformations. In §11, the scope of regular transformations is discussed briefly. Finally in §12 we indicate the possibility of taking point sets other than the set of positive numbers for the domains of s and t in kernel transformations.

2. **Properties of integral.** In the future we use those and only those definitions of integral having properties which we now specify. To simplify our statements of the properties, we use the symbol "$f \in I(a, b)$" to indicate that a and b are finite real numbers with $a<b$ and that $f(t)$ is a member of the class $I(a, b)$ of complex-valued functions integrable over the interval $a \leqq t \leqq b$.

I. *If $f_1, f_2 \in I(a, b)$ and c_1, c_2 are complex constants, then $c_1f_1+c_2f_2 \in I(a, b)$ and*

$$\int_a^b [c_1f_1(t) + c_2f_2(t)]dt = c_1\int_a^b f_1(t)dt + c_2\int_a^b f_2(t)dt. \tag{2.1}$$

II. *If f_1 and f_2 are real, then $f_1+if_2 \in I(a, b)$ if and only if $f_1 \in I(a, b)$ and $f_2 \in I(a, b)$.*

III. *If $f \in I(a, c)$ and $a<b<c$, then $f \in I(a, b)$, $f \in I(b, c)$, and*

$$\int_a^c f(t)dt = \int_a^b f(t)dt + \int_b^c f(t)dt. \tag{2.2}$$

If $a<b<c$, $f \in I(a, b)$, and $f \in I(b, c)$, then $f \in I(a, c)$ and (2.2) holds.

IV. *If* $a \leqq b$, *then*

$$\int_a^b 1\,dt = b - a. \tag{2.3}$$

V. *If* $f_1, f_2 \in I(a, b)$ *while* f_1 *and* f_2 *are real and* $f_1(t) \leqq f_2(t)$ *over* $a \leqq t \leqq b$, *then*

$$\int_a^b f_1(t)\,dt \leqq \int_a^b f_2(t)\,dt. \tag{2.4}$$

This implies that if $f \in I(a, b)$ *and* f *is real, then the integral of* f *is real.*

VI. *If* f, $|f| \in I(a, b)$, *then*

$$\left| \int_a^b f(t)\,dt \right| \leqq \int_a^b |f(t)|\,dt. \tag{2.5}$$

VII. *If, for each* h *with* $a < h < b$, f, $|f| \in I(a, h)$ *and*

$$\int_a^h |f(t)|\,dt < M, \tag{2.6}$$

M *being a constant independent of* h, *then* $f \in I(a, b)$ *and*

$$\int_a^b f(t)\,dt = \lim_{h \to b-} \int_a^h f(t)\,dt; \tag{2.7}$$

likewise, if $f \in I(h, b)$ *when* $a < h < b$ *and*

$$\int_h^b |f(t)|\,dt < M, \tag{2.8}$$

M *being a constant independent of* h, *then* $f \in I(a, b)$ *and*

$$\int_a^b f(t)\,dt = \lim_{h \to a+} \int_h^b f(t)\,dt; \tag{2.9}$$

except that, for any integral (and in particular for the Riemann integral) for which $f \in I(a, b)$ *implies that* $f(t)$ *must be bounded over* $a \leqq t \leqq b$, VII *shall apply only to functions* $f(t)$ *bounded over* $a \leqq t \leqq b$.

VIII. *The left member of*

$$\int_a^\infty f(t)\,dt = \lim_{h \to \infty} \int_a^h f(t)\,dt \tag{2.91}$$

is defined to be the right member whenever the right member exists.

Any integral satisfying I–VIII will be denoted hereafter by $\int$.

All of the conditions I to VII (that is, those which pertain to finite intervals) are satisfied by both Riemann and Lebesgue integrals. These integrals, extended over $(0, \infty)$ by (2.91), will be called simply Riemann and Lebesgue integrals respectively; hence Riemann and Lebesgue integrals satisfy I–VIII. Existence of

$$\int_a^\infty f(t)dt \tag{2.92}$$

does not imply existence of

$$\int_a^\infty |f(t)|\,dt. \tag{2.93}$$

However, properties I, III, V, VI, VIII, and the Cauchy criterion for convergence imply that, if $f \in I(a, h)$ for each $h > a$ and (2.93) exists, then (2.92) must exist.

Properties III, V, VI, and VII imply that $\int$ shares with Riemann and Lebesgue integrals the property of being independent of the value of the integrand at any finite set of points. For example, if $f \in I(a, b)$ and $g(t) = f(t)$ for $a < t < b$, then $g \in I(a, b)$ and

$$\int_a^b g(t)dt = \int_a^b f(t)dt$$

irrespective of whether $g(a)$ and $g(b)$ agree with $f(a)$ and $f(b)$ or are different from $f(a)$ and $f(b)$. Again, if $f \in I(a, b)$ and $g \in I(b, c)$ where $a < b < c$, and $h(x) = f(x)$ or $g(x)$ according as $a \leqq x \leqq b$ or $b < x \leqq c$, then $h \in I(a, c)$ and

$$\int_a^c h(x)dx = \int_a^b f(x)dx + \int_b^c g(x)dx.$$

Properties of $\int$ imply that, if f is real and bounded and $f \in I(a, b)$, then the integral over (a, b) of f lies in the closed interval bounded by the lower and upper Darboux (sometimes called Riemann) integrals over (a, b) of f.

3. **Definition of kernel.** We next decide what kind of functions $K(s, t)$ we shall admit as kernels in the transformation

$$K: \qquad y(s) = \int_0^\infty K(s, t)x(t)dt.$$

Suppose the integral chosen happens to be a Riemann integral. It is

then true that, if $x(t) \equiv 0$,

$$\int_0^h K(s, t)x(t)dt, \qquad h > 0, \tag{3.01}$$

will exist for every function $K(s, t)$; if, however, one wishes (3.01) to exist for all functions $x(t)$ in some significant class of functions, it would be at least unpleasant to have a function $K(s, t)$ which is not a Riemann integrable function (say unbounded or perhaps nonmeasurable) of t over $0 \leqq t \leqq h$ for each s. In other words, it seems reasonable to impose some "condition of integrability" on $K(s, t)$ to ensure that the transformation K will be significant. Accordingly we adopt the following definition.*

A complex-valued function $K(s, t)$ will be called a *kernel* if the integrals

$$\int_0^h K(s, t)dt, \qquad h, s > 0, \tag{3.02}$$

$$\int_0^h K(s_1, t) \operatorname{sgn} K(s_2, t)dt, \qquad h, s_1, s_2 > 0, \tag{3.03}$$

$$\int_0^h K(s_1, t)\Re \operatorname{sgn} K(s_2, t)dt, \qquad h, s_1, s_2 > 0, \tag{3.04}$$

all exist. Existence of (3.03) and (3.04) implies existence of

$$\int_0^h K(s_1, t)\Im \operatorname{sgn} K(s_2, t)dt, \qquad h, s_1, s_2 > 0. \tag{3.05}$$

Setting $s_1 = s_2 = s$ in (3.03) gives existence of

$$\int_0^h |K(s, t)|\,dt, \qquad h, s > 0. \tag{3.06}$$

The class of functions $K(s, t)$ satisfying (3.02), (3.03), and (3.04) is the class referred to in §1. Accordingly K denotes in this paper either a function satisfying (3.02), (3.03), and (3.04), or the transformation determined by such a kernel.

For the case of Lebesgue integrals, existence of (3.03) and (3.04) need not be explicitly required since it is implied by existence of (3.02). For the case of Riemann integrals, however, existence of (3.02) does not imply that of (3.03) and (3.04). The latter fact is readily

* If u and v are real and $z = u + iv$, then $\Re z = u$, $\Im z = v$; and sgn z is 0 or $|z|/z$ according as $z = 0$ or $z \neq 0$, so that in every case $z \operatorname{sgn} z = |z|$.

proved by use of the function $g(t)$ defined as follows. Let $r_1, r_2, \cdots$ denote in some order the positive rational numbers, and let $g(t) = r_n/n$ or 0 according as $t = r_n$ or t is irrational. In each interval $(0, h)$, $g(t)$ has the Riemann integral 0 while sgn $g(t)$ is not Riemann integrable. Thus we have here one more place in analysis where a theory based on use of Lebesgue integrals avoids assumptions [in this case (3.03) and (3.04)] which are required for a corresponding theory based on Riemann integrals.

The conditions (3.02), (3.03), and (3.04) do not (unless $\int$ happens to be of a restrictive type) imply that $K(s, t)$ is bounded over each interval $0 < t < h$. For example, if $\Re r \neq 1$ the function (1.04) associated with the Hölder transformation (1.03) of order r, being unbounded over $0 < t < s$ for $s > 0$, would not be a kernel when $\int$ is the proper Riemann integral; but when $\int$ is either the Cauchy-Riemann improper integral or the Lebesgue integral, the function is a kernel when $\Re r > 0$.

4. **The class $\mathfrak{K}$.** If a function x is such that

$$y(s) = \int_0^\infty K(s, t)x(t)dt \tag{4.01}$$

exists for each $s > 0$, then conditions VIII and III of §2 imply that

$$F(h, s) = \int_0^h K(s, t)x(t)dt \tag{4.02}$$

must exist for each h, $s > 0$. Let $\mathfrak{K}$ be the class of functions x for which $F(h, s)$ exists for h, $s > 0$. The condition $x \in \mathfrak{K}$ is essentially a local "condition for integrability," and is not concerned with the behavior of $x(t)$ as $t \to \infty$. If the kernel $K(s, t)$ is of *finite reference* (that is, if for each s, $K(s, t) = 0$ for all sufficiently great t), then $x \in \mathfrak{K}$ implies existence of the K transform $y(s)$; but otherwise $x \in \mathfrak{K}$ does not ordinarily imply existence of $y(s)$. For many transformations, the condition $x \in \mathfrak{K}$ is precisely the condition that $x(t)$ be integrable over each finite interval $(0, h)$.

The class $\mathfrak{K}$ is linear, that is, if $x_1, x_2 \in \mathfrak{K}$ and c_1, c_2 are complex constants, then $c_1x_1 + c_2x_2 \in \mathfrak{K}$.

The conditions which we have imposed upon $\int$ and K are sufficient to make $\mathfrak{K}$ an extensive class of functions. Indeed, some of the conditions were imposed to ensure that $\mathfrak{K}$ contains the functions $x_0(t)$ which are described in the following lemma, and which we shall use later.

LEMMA 4.1. *Let* $0=h_1<h_2<\cdots$ *be a sequence for which* $\lim_{n\to\infty}h_n=\xi$ *may be either finite or* $+\infty$. *Let* $s_1, s_2, \cdots$ *be a sequence of positive numbers. Let* $\sigma_1, \sigma_2, \cdots$ *be a sequence of real numbers with* $|\sigma_n| \leqq 1$ *for each* n. *Let* $x_0(t)$ *be defined by the formulas*

$$x_0(t) = \sigma_n \operatorname{sgn} K(s_n, t), \qquad h_n \leqq t < h_{n+1}, \tag{4.11}$$

and, in case $\xi<\infty$, *by the additional formula*

$$x_0(t) = 0, \qquad t \geqq \xi. \tag{4.12}$$

Then $x_0 \in K$, $\Re x_0 \in K$, *and* $\Im x_0 \in K$.

It is easy to modify Lemma 4.1 and its proof to cover functions $x_0(t)$ defined analogously to (4.11) over intervals whose end points form a decreasing rather than an increasing sequence.

To prove that $\Re x_0 \in K$, let $s>0$ and $h>0$ be fixed. Our hypotheses on f and K imply existence of

$$\int_{h_n}^{h_{n+1}} K(s, t)\Re \operatorname{sgn} K(s_n, t)dt, \qquad n = 1, 2, \cdots. \tag{4.13}$$

Since σ_n is real, it follows that, if we set for simplicity $x_1(t)=\Re x_0(t)$, then

$$\int_{h_n}^{h_{n+1}} K(s, t)x_1(t)dt, \qquad n = 1, 2, \cdots, \tag{4.14}$$

exists. Therefore, for each $m=1, 2, \cdots$,

$$F_1(b, s) = \int_0^b K(s, t)x_1(t)dt \tag{4.15}$$

exists when $b=h_m$. If $h<\xi$, we can choose $h_m>h$ and conclude existence of $F_1(h, s)$. If $\xi=+\infty$, this implies that $x_1=\Re x_0 \in K$. If $\xi<\infty$, then existence of $F_1(b, s)$ for $b<\xi$, the fact that $K(s, t)x_1(t)$ is bounded over $0\leqq t\leqq\xi$ if $K(s, t) \in I(0, \xi)$ requires that $K(s, t)$ be bounded over $0\leqq t\leqq\xi$, and the inequality

$$\int_0^b |K(s, t)x_1(t)|\,dt \leqq \int_0^\xi |K(s, t)|\,dt, \qquad 0 < b < \xi, \tag{4.16}$$

imply by VII of §2 that $F_1(\xi, s)$ exists. If $h>\xi$, then existence of $F_1(\xi, s)$ and (4.12) imply existence of $F_1(h, s)$. Therefore $\Re x_0=x_1 \in K$. We can prove in the same way that $\Im x_0 \in K$, and linearity of K then gives $x_0 \in K$.

5. **Regularity.** We recall that a matrix transformation A is said to be regular if each convergent sequence is summable A to the value to which it converges, that is, if existence of $\lim x_t$ implies existence of $y_0, y_1, \cdots$ and the equality $\lim y_s = \lim x_t$. Necessary and sufficient conditions for regularity of A are, by the Silverman-Toeplitz theorem,

$$\sum_{t=0}^{\infty} |A_{st}| < M, \qquad s = 0, 1, 2, \cdots, \tag{5.01}$$

$$\lim_{s\to\infty} A_{st} = 0, \qquad t = 0, 1, 2, \cdots, \tag{5.02}$$

$$\lim_{s\to\infty} \sum_{t=0}^{\infty} A_{st} = 1, \tag{5.03}$$

M being a constant independent of s.

We use this theorem, and the theorems given later in this paragraph, merely for purposes of analogy; their truth is well known and will be demonstrated at the end of §10. The matrix A (or matrix transformation A) is *regular over a class C* of sequences if each convergent sequence $x \in C$ is summable A to the value to which it converges. The conditions (5.01) and (5.02) are necessary and sufficient to ensure that A be regular over the class of null sequences (sequences converging to 0). Also the conditions (5.01), (5.02), and (5.03) are necessary as well as sufficient for regularity of A (which may or may not be real) over the class of real sequences.

If we should call a kernel transformation K regular only when it has the property that each function $x(t)$ for which $\lim_{t\to\infty} x(t)$ exists has a transform $y(s)$ for which $\lim_{s\to\infty} y(s) = \lim_{t\to\infty} x(t)$, then it would turn out that no kernel transformation whatever involving a Riemann or Lebesgue integral could be regular. Consider, for example, the transformation

$$y(s) = \frac{1}{s}\int_0^s x(t)dt, \qquad s > 0, \tag{5.04}$$

in which the integral is that of Lebesgue. If $x(t) = 1/t^2$, then $\lim x(t) = 0$; but $y(s)$ does not exist (may be said to be $+\infty$) for each $s > 0$ and accordingly $\lim_{s\to\infty} y(s) = 0$ fails. Also, if $x(t)$ is a bounded function for which $\lim x(t) = 0$ but which is non-measurable over each finite interval, then again $y(s)$ fails to exist for each $s > 0$ and $\lim_{s\to\infty} y(s) = 0$ fails. However, if $x(t)$ belongs to the class of functions which are integrable over each finite interval and converge

as $t \to \infty$, then it is well known and easy to show that $y(s)$ defined by (5.04) exists and that $\lim_{s\to\infty} y(s) = \lim_{t\to\infty} x(t)$.

These considerations imply that the following definition, which is analogous to one involving matrix transformations, will be useful. The transformation K is *regular over the class* C of functions x if each $x \in C$ for which $\lim x(t)$ exists is summable K to $\lim x(t)$. For example, (5.04) with Lebesgue integral is regular over the class C_1 of continuous functions and is also regular over the larger class C_2 of functions Lebesgue integrable over each finite interval $(0, h)$.

The largest class of functions over which a transformation K could possibly be regular is the class K_∞ of §4. Accordingly, we define K to be *regular* if it is regular over the class K_∞.

Regularity is only one of several properties in which we shall be interested. However, the conditions which characterize regular transformations K are so important in the theory of summability that it seems desirable to present them here and to discuss them briefly.

THEOREM 5.1. *In order that K be regular over K_∞, it is necessary and sufficient that*

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > 0, \tag{5.11}$$

$$\limsup_{s\to\infty} \int_0^\infty |K(s,t)|\,dt = M < \infty, \tag{5.12}$$

$$\lim_{s\to\infty} \int_0^h K(s,t)x(t)dt = 0, \qquad h > 0,\ x \in K_\infty, \tag{5.13}$$

$$\lim_{s\to\infty} \int_0^\infty K(s,t)dt = 1. \tag{5.14}$$

In the statement of this theorem, the phrase "K be regular over K_∞" is used instead of the equivalent phrase "K be regular" in order to facilitate the statement of closely related theorems in §9. Theorem 5.1 will be proved in §9. Meanwhile we assume it and discuss the conditions involved.

The analogy between (5.01) on the one hand and (5.11) and (5.12) on the other hand seems satisfactory when we bear in mind that a sequence $\Sigma_0, \Sigma_1, \Sigma_2, \cdots$ of real numbers is bounded whenever $\limsup |\Sigma_n| < \infty$, but that a corresponding conclusion involving functions cannot be drawn. It is obvious either from the definition of regularity or from the conditions of Theorem 5.1 that, if $K_1(s,t)$ is a regular kernel, and $\phi(s)$ is any function of s which is defined for

$s>0$ and converges to 1 as $s\to\infty$ (say $\phi(s)=1+\psi(s)/s^2$ where $\psi(s)$ is a bounded non-measurable function which converges to 1 as $s\to 0$), then $K_2(s, t)=\phi(s)K_1(s, t)$ is also regular. In particular, if $K_1(s, t)$ is the simple kernel of (1.02), then both K_1 and K_2 are regular, and the equality

$$\int_0^\infty |K_2(s, t)|\,dt = \left|\int_0^\infty K_2(s, t)dt\right| = |\phi(s)| \tag{5.15}$$

shows that the condition

$$\text{l.u.b.}_{s>0} \int_0^\infty |K(s, t)|\,dt < \infty, \tag{5.16}$$

which is analogous to (5.01), is not necessary for regularity of K. Indeed, the choice of K_1 and ϕ suggested above shows that K can be regular even though

$$\text{l.u.b.}_{s>0} \int_0^h |K(s, t)|\,dt = \infty, \qquad h>0. \tag{5.17}$$

The condition (5.14) is an exact analogue of (5.03). The condition (5.13) is by no means as attractive as the corresponding condition (5.02). However the real test of (5.13) comes when one seeks to determine whether a given kernel satisfies it. Actually it is often easier to determine whether a given kernel satisfies (5.13) than to determine whether it satisfies (5.12), and from this point of view we may accept (5.13). We do not in this paper make any attempt to present (5.13) in a different or more attractive form. Some consequences of the conditions for regularity are given by the following theorem.

THEOREM 5.2. *If K is regular, then*

$$\lim_{s\to\infty} \int_0^h K(s, t)dt = 0, \qquad h>0, \tag{5.21}$$

$$\lim_{s\to\infty} \int_h^\infty K(s, t)dt = 1, \qquad h\geqq 0, \tag{5.22}$$

and

$$1 \leqq \lim_{h\to\infty} \limsup_{s\to\infty} \int_h^\infty |K(s, t)|\,dt \leqq M < \infty. \tag{5.23}$$

The condition (5.21) is obtained by setting $x(t)=1$ in (5.13); and (5.22) is implied by (5.14) and (5.12). The inequality

$$1 \leqq \limsup_{s\to\infty} \int_h^\infty |K(s,t)|\,dt \leqq M, \qquad h \geqq 0, \tag{5.24}$$

is implied by (5.12) and (5.22). Since the middle term of (5.24) is a monotone decreasing function of h, (5.24) implies (5.23).

The number

$$\Lambda = \lim_{h\to\infty} \limsup_{s\to\infty} \int_h^\infty |K(s,t)|\,dt \tag{5.25}$$

of condition (5.23) turns out to be significant in determining the properties which K has or fails to have. If Λ is defined by (5.25), then

$$\Lambda = \limsup_{h,s\to\infty} \int_h^\infty |K(s,t)|\,dt; \tag{5.26}$$

and conversely if Λ is defined by (5.26), then (5.25) holds.

It can be shown by an example that (5.11), (5.12), (5.21), and (5.14) are not sufficient to ensure regularity of K.

The condition $\lim_{s\to\infty} A_{st} = 0$ being necessary for regularity of a matrix A, so also is the condition

$$\lim_{s\to\infty} \sum_{t=0}^{h} |A_{st}| = 0, \qquad h = 0, 1, 2, \cdots. \tag{5.27}$$

To emphasize the fact that conditions analogous to these are not necessary for regularity of K, we give

THEOREM 5.3. *Neither of the conditions*

$$\lim_{s\to\infty} K(s,t) = 0, \qquad t > 0, \tag{5.31}$$

$$\lim_{s\to\infty} \int_0^h |K(s,t)|\,dt = 0, \qquad h > 0, \tag{5.32}$$

is necessary for regularity of K.

We prove this theorem by an example in which the integral is that of Lebesgue, and $x \in \mathcal{K}$ is equivalent to the condition that $x(t)$ be Lebesgue integrable over each finite interval $(0, h)$. If $x \in \mathcal{K}$ and $\limsup |x(t)| < \infty$, then

$$\begin{aligned} y(s) = &\int_0^s \frac{e^{ist}}{1+t^2}\,x(t)dt + \int_s^{s+1} x(t)dt \\ &+ \int_{s+1}^\infty \frac{e^{is(t-1)}}{1+(t-1)^2}\,x(t)dt \end{aligned} \tag{5.33}$$

exists for each $s \geqq 0$. For each $s \geqq 0$, $K(s, t)$ is $e^{ist}/(1+t^2)$ or 1 or $e^{is(t-1)}/[1+(t-1)^2]$ according as $0 \leqq t < s$ or $s \leqq t < s+1$ or $s+1 \leqq t$. Elementary integration gives, for each $s \geqq 0$,

$$\int_0^\infty |K(s, t)|\, dt = \int_0^\infty \frac{1}{1+t^2}\, dt + \int_s^{s+1} 1 ds = \frac{\pi}{2} + 1. \tag{5.34}$$

By the well known Riemann-Lebesgue theorem

$$\lim_{s\to\infty} \int_0^h K(s, t)x(t)dt = \lim_{s\to\infty} \int_0^h e^{ist}[x(t)/(1+t^2)]dt = 0 \tag{5.35}$$

for $h>0$, $x \in \mathfrak{K}$. Also,

$$\int_0^\infty K(s, t)dt = 1 + \int_0^\infty e^{ist}[1/(1+t^2)]dt,$$

and another application of the Riemann-Lebesgue theorem gives

$$\lim_{s\to\infty} \int_0^\infty K(s, t)dt = 1. \tag{5.36}$$

It follows from Theorem 5.1 that the transformation (5.33) is regular. But there is no $t>0$ for which $\lim_{s\to\infty} K(s, t)$ exists; and if $0 \leqq a < b$, then for all sufficiently great s

$$\int_a^b |K(s, t)|\, dt = \int_a^b \frac{1}{1+t^2}\, dt > 0,$$

so that

$$\lim_{s\to\infty} \int_a^b |K(s, t)|\, dt \tag{5.37}$$

exists and is positive. This proves Theorem 5.3.

We give an example of a simple pair of regular transformations of finite reference to show that the conditions (5.31) and (5.32) are not significant in the theory of summability. Let K_1 and K_2 be defined by

$$K_1: \qquad y_1(s) = \int_0^1 \sigma(s)e^{ist}x(t)dt + \int_s^{s+1} x(t)dt,$$

$$K_2: \qquad y_2(s) = \int_s^{s+1} x(t)dt,$$

where $\sigma(s)$ is 0 or 1 according as $[s]$, the greatest integer less than or equal to s, is even or odd, and where the integrals are Lebesgue integrals. For K_1, the limits in (5.31) and (5.32) both fail to exist; for

K_2, both (5.31) and (5.32) hold. But $K_{\mathcal{L}1} \equiv K_{\mathcal{L}2}$, each being the class of functions integrable over each finite interval $(0, h)$; and if $x \in K_{\mathcal{L}1}$, then by the Riemann-Lebesgue theorem,

$$\lim_{s\to\infty} | y_2(s) - y_1(s) | = 0.$$

Hence K_1 and K_2 are, as methods of summability, substantially identical.

In the next two sections, we give theorems which show that even very moderate hypotheses on a transformation K imply that (5.11) and (5.12) must hold.

6. **Application of K to functions $x \in \mathcal{R}K_{\mathcal{L}B}(h)$.** Let $K_{\mathcal{L}B}(h)$ denote the class of functions $x(t)$ such that $x \in K_{\mathcal{L}}$, $x(t)$ is bounded, and $x(t) = 0$ for all $t > h$. If C is a class of functions, we use $\mathcal{R}C$ to denote the subclass of C consisting of the real functions in C. In several theorems of this and the next section, H denotes a real nonnegative constant which will be 0 in most of our applications.

THEOREM 6.1. *Let $h > 0$ and $H \geqq 0$ be fixed. If K is such that, for each $x \in \mathcal{R}K_{\mathcal{L}B}(h)$,*

$$y(s) = \int_0^\infty K(s, t)x(t)dt \equiv \int_0^h K(s, t)x(t)dt \tag{6.11}$$

exists for $s > H$ and has the property

$$\text{l.u.b.}_{s>H} | y(s) | < \infty, \tag{6.12}$$

then there is a constant $M \equiv M(h) < \infty$ such that

$$\text{l.u.b.}_{s>H} \int_0^h | K(s, t) | dt = M. \tag{6.13}$$

THEOREM 6.2. *Let $h > 0$ be fixed. If K is such that, for each $x \in \mathcal{R}K_{\mathcal{L}B}(h)$,*

$$y(s) = \int_0^\infty K(s, t)x(t)dt \equiv \int_0^h K(s, t)x(t)dt \tag{6.21}$$

exists for $s > s_0(x)$, where $s_0(x)$ is a real number for each $x \in \mathcal{R}K_{\mathcal{L}B}(h)$, and has the property

$$\limsup_{s\to\infty} | y(s) | < \infty, \tag{6.22}$$

then there is a constant $M \equiv M(h) < \infty$ such that

$$\limsup_{s\to\infty} \int_0^h |K(s,t)|\,dt = M. \tag{6.23}$$

We first prove Theorem 6.1 by assuming that

$$\text{l.u.b.}_{s>H} \int_0^h |K(s,t)|\,dt = +\infty, \tag{6.24}$$

and constructing a real function $x \in K$ such that $|x(t)| \leqq 1$, $x(t)=0$ for $t>h$, and the transform $y(s)$ of x exists for $s>0$ and has the property $\text{l.u.b.}|y(s)| = +\infty$.

Let $a_0=0$, $b_0=h$. When a_n and b_n have been chosen such that

$$\text{l.u.b.}_{s>H} \int_{a_n}^{b_n} |K(s,t)|\,dt = +\infty, \tag{6.25}$$

the properties of $\int$ given in §2 imply that at least one of the two definitions $[a_{n+1}=a_n;\ b_{n+1}=(a_n+b_n)/2]$ and $[a_{n+1}=(a_n+b_n)/2;\ b_{n+1}=b_n]$ will make (6.25) hold when n is replaced by $(n+1)$. Thus we obtain by induction a monotone increasing sequence $a_0, a_1, a_2, \cdots$ and a monotone decreasing sequence $b_0, b_1, b_2, \cdots$, such that $b_n-a_n=h/2^n$ and (6.25) holds for each $n=0, 1, 2, \cdots$. Let ξ be the common limit of the sequences a_n and b_n, so that

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \xi. \tag{6.26}$$

It is easy to show that at least one of the two functions of s

$$\int_{a_n}^{\xi} |K(s,t)|\,dt, \qquad \int_{\xi}^{b_n} |K(s,t)|\,dt, \tag{6.27}$$

is unbounded for an infinite set of n and hence for each $n=0, 1, 2, \cdots$. Treatment of the alternative case being similar, we consider only the case that the first is unbounded for each $n=0, 1, 2, \cdots$. In this case

$$F(s,c) = \int_c^{\xi} |K(s,t)|\,dt, \qquad s > H, \tag{6.28}$$

is an unbounded function of s for each c in the interval $0 \leqq c < \xi$.

Let $c_1=0$. Since $F(s, c_1)$ is unbounded, we can choose $s_1>H$ such that $F(s_1, c_1)>2+2^{-1}$. Since existence of (3.06) implies that $\lim_{c\to\xi} F(s_1, c)=0$, we can choose c_2 such that $c_1<c_2<\xi$, $\xi-c_2<2^{-2}$, and $F(s_1, c_2)<2^{-1}$. Next, choose $s_2>H$ such that $F(s_2, c_2)>2^2+2^{-2}$, and then choose c_3 such that $c_2<c_3<\xi$, $\xi-c_3<2^{-3}$, and $F(s_2, c_3)<2^{-2}$. Proceeding in this manner, we obtain a sequence $s_1, s_2, \cdots$ of values

of $s>H$ and a sequence $c_1<c_2<c_3<\cdots$ of values of c such that $c_n\to\xi$, and for each $n=1, 2, \cdots$

$$\int_{c_n}^{\xi} |K(s_n, t)|\,dt > 2^n + 2^{-n}, \quad \int_{c_{n+1}}^{\xi} |K(s_n, t)|\,dt < 2^{-n}, \tag{6.29}$$

and therefore also

$$\int_{c_n}^{c_{n+1}} |K(s_n, t)|\,dt > 2^n. \tag{6.31}$$

We are now ready to define a function $x_0(t)$ in terms of which the desired function $x(t)$ will be determined. The function x_0 will be of the type described in Lemma 4.1, and accordingly $x_0 \in K$. First, let

$$x_0(t) = \operatorname{sgn} K(s_1, t), \qquad 0 = c_1 \leqq t < c_2. \tag{6.32}$$

Then, using (6.31) with $n=1$, we find that

$$\left|\int_0^{c_{n+1}} K(s_n, t)x_0(t)dt\right| > 2^{n-1} \tag{6.33}$$

holds when $n=1$. When we have extended the definition of $x_0(t)$ over the interval $0\leqq t<c_{n+1}$ in such a way that (6.33) holds, we define $x_0(t)$ for $c_{n+1}\leqq t<c_{n+2}$ by the formula

$$x_0(t) = \sigma_{n+1} \operatorname{sgn} K(s_{n+1}, t), \tag{6.34}$$

where σ_{n+1} is 0 if

$$\left|\int_0^{c_{n+1}} K(s_{n+1}, t)x_0(t)dt\right| > 2^n \tag{6.35}$$

holds, and σ_{n+1} is 1 if (6.35) fails. If (6.35) holds, then $\sigma_{n+1}=0$, so that

$$\left|\int_0^{c_{n+2}} K(s_{n+1}, t)x_0(t)dt\right| = \left|\int_0^{c_{n+1}} K(s_{n+1}, t)x_0(t)dt\right| > 2^n. \tag{6.36}$$

If (6.35) fails, then $\sigma_{n+1}=1$ and our inequalities give

$$\begin{aligned}\left|\int_0^{c_{n+2}} K(s_{n+1}, t)x_0(t)dt\right| &\geqq \left|\int_{c_{n+1}}^{c_{n+2}} K(s_{n+1}, t)x_0(t)dt\right| - \left|\int_0^{c_{n+1}} K(s_{n+1}, t)x_0(t)dt\right| \\ &\geqq \int_{c_{n+1}}^{c_{n+2}} |K(s_{n+1}, t)|\,dt - 2^n \geqq 2^{n+1} - 2^n = 2^n.\end{aligned}$$

Thus, in both of the two cases, we find that (6.33) holds when n is replaced by $n+1$.

This induction serves to define $x_0(t)$ over $0 \leqq t < \xi$; let $x_0(t) = 0$ for $t \geqq \xi$. Then, since $x_0 \in K_B(h)$, the transform $y_0(s)$ of x_0 exists for each $s > 0$. For each $n = 1, 2, \cdots$ we find

$$|y_0(s_n)| = \left| \int_0^{\xi} K(s_n, t) x_0(t) dt \right|$$

$$\geqq \left| \int_0^{c_{n+1}} K(s_n, t) x_0(t) dt \right| - \int_{c_{n+1}}^{\xi} |K(s_n, t) dt| > 2^{n-1} - 2^{-n}.$$

This implies that l.u.b. $|y_0(s)| = +\infty$. While $x_0 \in K_B(h)$, it is not true that x_0 need be real unless K is real. Let $x_0(t) = x_1(t) + ix_2(t)$, where $x_1(t)$ and $x_2(t)$ are real. Then by Lemma 4.1, x_1, $x_2 \in K$ and hence $x_1, x_2 \in K_B(h)$. If we let $y_1(s)$ and $y_2(s)$ denote the transforms of x_1 and x_2, then

$$|y_1(s_n) + iy_2(s_n)| = |y_0(s_n)| > 2^{n-1} - 2^{-n}. \tag{6.37}$$

This implies that not both l.u.b. $y_1(s)$ and l.u.b. $y_2(s)$ can be finite. Accordingly, if $x(s)$ denotes a properly chosen one of the functions $x_1(s)$ and $x_2(s)$, then $x \in \mathcal{R}K_B(h)$ and the transform y of x is unbounded. This proves Theorem 6.1.

Proof of Theorem 6.2 is practically identical with that of Theorem 6.1. Contradiction of the conclusion (6.23), which is weaker than (6.13), gives

$$\limsup_{s \to \infty} \int_0^h |K(s, t)| dt = +\infty, \tag{6.38}$$

which is stronger than (6.24) and enables us to choose the sequence $s_1, s_2, \cdots$ in such a way that $s_{n+1} > s_n + 1$ and hence $s_n \to \infty$ as $n \to \infty$. Then (6.37) would imply lim sup $|y(s)| = +\infty$ and thus complete the proof of Theorem 6.2.

7. **Application of K to functions $x \in K_0$.** Let K_0 denote the subclass of functions $x \in K$ which are bounded over $0 < t < \infty$ and converge to 0 as $t \to \infty$. It is clear that for each $h > 0$ the class $K_B(h)$ is a subclass of K_0. The main theorems which we prove in this section are 7.4 and 7.5. We give first four simpler theorems which we shall need.

THEOREM 7.1. *If K is such that for each $x \in \mathcal{R}K_0$*

$$\limsup_{h \to \infty} \left| \int_0^h K(s, t) x(t) dt \right| < \infty, \qquad s > H, \tag{7.11}$$

then

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > H. \tag{7.12}$$

If (7.12) fails, then for some $s_0 > H$,

$$\lim_{h\to\infty} \int_0^h |K(s_0,t)|\,dt = \infty,$$

and we can choose a sequence $0 = h_1 < h_2 < \cdots$ of integers such that

$$\int_{h_n}^{h_{n+1}} |K(s_0,t)|\,dt \geqq n, \qquad n = 1, 2, \cdots.$$

If we set, for each $n = 1, 2, \cdots$,

$$x_0(t) = (1/n)\ \operatorname{sgn} K(s_0,t), \qquad h_n \leqq t < h_{n+1},$$

we find, using Lemma 4.1, that $x_0 \in K_0$ and

$$\int_0^{h_{m+1}} K(s_0,t)x_0(t)dt = \sum_{n=1}^{m} \frac{1}{n} \int_{h_n}^{h_{n+1}} |K(s,t)|\,dt \geqq m, \qquad m = 1, 2, \cdots. \tag{7.13}$$

Hence

$$\limsup_{h\to\infty} \left| \int_0^h K(s_0,t)x_0(t)dt \right| = \infty. \tag{7.14}$$

Using again Lemma 4.1, we see that $x_1 = \mathcal{R}x_0$ and $x_2 = 3x_0$ are both members of $\mathcal{R}K_0$. It then follows from (7.14) that the statement obtained by replacing x_0 in (7.14) by one or the other of x_1 or x_2 must be true. This contradicts the hypothesis of Theorem 7.1, and Theorem 7.1 is accordingly proved. This proof has been so phrased that it serves also as a proof of the following theorem.

THEOREM 7.15. *If K is such that $K(s,t)$ is independent of t over each interval $n \leqq t < n+1$, $n = 0, 1, 2, \cdots$, and if*

$$\limsup_{n\to\infty} \left| \int_0^n K(s,t)x(t)dt \right| < \infty, \qquad s > H, \tag{7.16}$$

for each real function $x(t)$ which is constant over each interval $n \leqq t < n+1$ and converges to 0 as $t \to \infty$, then

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > H. \tag{7.17}$$

THEOREM 7.2. *If K is such that for each $x \in \mathcal{R}K_0$*

$$\limsup_{h\to\infty} \left| \int_0^h K(s,t)x(t)dt \right| < \infty \tag{7.21}$$

for $s > s_0(x)$, where $s_0(x) < \infty$ is a real number for each $x \in \mathcal{R}K_0$, then a constant $H < \infty$ exists such that

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > H. \tag{7.22}$$

This theorem would reduce essentially to Theorem 7.1 if it were assumed that a number $H < \infty$ exists such that $s_0(x) < H$ for all $x \in \mathcal{R}K_0$; but this is not assumed and accordingly Theorem 7.2 is essentially different from Theorem 7.1. To prove Theorem 7.2, we assume that the conclusion fails and obtain a contradiction of the hypothesis. Since

$$\int_0^h |K(s,t)|\,dt$$

exists for h, $s > 0$, failure of the conclusion implies existence of a sequence $0 < s_1 < s_2 < s_3 < \cdots$ such that $s_n \to \infty$ and

$$\int_0^\infty |K(s_n,t)|\,dt = \lim_{h\to\infty} \int_0^h |K(s_n,t)|\,dt = \infty, \quad n = 1,2,\cdots. \tag{7.23}$$

Let $h_0 = 0$ and choose h_1 such that

$$\int_{h_0}^{h_1} |K(s_1,t)|\,dt > 2.$$

Next, choose $h_2 > h_1 + 1$ such that

$$\int_{h_1}^{h_2} |K(s_p,t)|\,dt > 2^2, \qquad p = 1,2.$$

When the $h_0, h_1, \cdots, h_{n-1}$ are determined, choose $h_n > h_{n-1} + 1$ such that

$$\int_{h_{n-1}}^{h_n} |K(s_p,t)|\,dt > 2^n, \qquad p = 1,2,\cdots,n. \tag{7.24}$$

Let $\alpha_1, \alpha_2, \cdots$ denote in order the integers in the sequence

$$1, 1, 2, 1, 2, 3, 1, 2, 3, 4, \cdots, 1, 2, \cdots, m, \cdots.$$

Since $\alpha_n \leqq n$ for each $n = 1, 2, \cdots$, it follows from (7.24) that

$$\int_{h_{n-1}}^{h_n} |K(s_{\alpha_n}, t)|\, dt > 2^n, \qquad n = 1, 2, \cdots. \tag{7.25}$$

Let $x_0(t)$ be defined over $0 \leqq t < \infty$ by the formulas

$$x_0(t) = n^{-1} \operatorname{sgn} K(s_{\alpha_n}, t), \qquad h_{n-1} \leqq t < h_n. \tag{7.26}$$

Then, by Theorem 4.1, $x_0 \in K_c$, and (7.26) implies further that $x_0 \in K_{c0}$. Now let q be a positive integer, and let

$$F_q(h) = \int_0^h K(s_q, t) x_0(t) dt. \tag{7.27}$$

For each of the infinite set of values of n for which $\alpha_n = q$, we can use (7.27), (7.26) and (7.25) to obtain

$$\begin{aligned} F_q(h_n) - F_q(h_{n-1}) &= \int_{h_{n-1}}^{h_n} K(s_{\alpha_n}, t) x_0(t) dt \\ &= \frac{1}{n} \int_{h_{n-1}}^{h_n} |K(s_{\alpha_n}, t)|\, dt > n^{-1} 2^n. \end{aligned}$$

Therefore,

$$\limsup_{h \to \infty} \left| \int_0^h K(s_q, t) x_0(t) dt \right| = \infty \tag{7.28}$$

for each $q = 1, 2, 3, \cdots$. If we let $x_1 = \Re x_0$, $x_2 = \Im x_0$, then $x_1, x_2 \in K_c$ by Theorem 4.1 and hence $x_1, x_2 \in \mathcal{R}K_{c0}$. Moreover, if x_0 is replaced in (7.28) by one or the other (perhaps either) of the functions x_1 and x_2, then (7.28) must hold for an infinite set of values of q. Since $s_q \to \infty$, this contradicts the hypothesis of Theorem 7.2, and the proof of Theorem 7.2 is complete.

THEOREM 7.3. *Let $H \geqq 0$ be fixed. If*

$$\int_0^\infty |K(s, t)|\, dt < \infty, \qquad s > H, \tag{7.31}$$

then

$$y(s) = \int_0^\infty K(s, t) x(t) dt, \qquad s > H, \tag{7.32}$$

exists for each $x \in K_c$ such that $\limsup |x(t)| < \infty$.

Let $x \in K_c$, $\limsup |x(t)| < \infty$, and choose C and $t_0 > 0$ such that $|x(t)| < C$ when $t > t_0$. Let $s > H$ be fixed. Then, since $x \in K_c$, it follows that $F(h) = \int_0^h K(s, t) x(t) dt$ exists for each $h > 0$, and if $q > p > t_0$ then

$$| F(q) - F(p) | = \left| \int_p^q K(s, t)x(t)dt \right| \leqq C \int_p^q | K(s, t) | \, dt.$$

This inequality and the Cauchy criterion for convergence imply existence of $y(s) = \lim F(h)$, and Theorem 7.3 is proved.

THEOREM 7.4. *Let $H \geqq 0$ be fixed. If K is such that for each $x \in \mathcal{RK}_0$*

$$y(s) = \int_0^\infty K(s, t)x(t)dt, \qquad s > H, \tag{7.41}$$

exists and has the property

$$\text{l.u.b.}_{s>H} \, | y(s) | < \infty, \tag{7.42}$$

then there is a constant $M < \infty$ such that

$$\text{l.u.b.}_{s>H} \int_0^\infty | K(s, t) | \, dt = M. \tag{7.43}$$

THEOREM 7.5. *If K is such that, for each $x \in \mathcal{RK}_0$,*

$$y(s) = \int_0^\infty K(s, t)x(t)dt \tag{7.51}$$

exists for $s \geqq s_0(x)$, where $s_0(x) < \infty$ is a real number for each $x \in \mathcal{RK}_0$, and has the property

$$\limsup_{s\to\infty} | y(s) | < \infty, \tag{7.52}$$

then there is a constant $M < \infty$ such that

$$\limsup_{s\to\infty} \int_0^\infty | K(s, t) | \, dt = M. \tag{7.53}$$

We first prove Theorem 7.4. The hypothesis and Theorem 7.1 imply that

$$\int_0^\infty | K(s, t) | \, dt < \infty, \qquad s > H. \tag{7.61}$$

Also, since $\mathcal{RK}_B(h)$ is a subclass of $\mathcal{RK}_0$, it follows from our hypothesis and Theorem 6.1 that for each $h > 0$ there is a constant $M(h)$ such that

$$\text{l.u.b.}_{s>H} \int_0^h |K(s,t)|\,dt = M(h). \tag{7.62}$$

Assuming that (7.43) fails, that is, that

$$\text{l.u.b.}_{s>H} \int_0^\infty |K(s,t)|\,dt = \infty, \tag{7.63}$$

we construct a function $x \in \mathcal{R}K_0$ for which (7.42) fails. From (7.62) and (7.63), it follows that

$$\text{l.u.b.}_{s>H} \int_h^\infty |K(s,t)|\,dt = \infty, \qquad h \geqq 0. \tag{7.64}$$

Let $h_0=0$ and $M(h_0)=0$. It follows from (7.64) that we can choose $s_1>H$ such that the inequality

$$\int_{h_{n-1}}^\infty |K(s_n,t)|\,dt > 4^n[2+M(h_{n-1})] \tag{7.65}$$

will hold when $n=1$, and then choose h_1 so that $h_1>h_0+1$ and

$$\begin{aligned} &\int_{h_{n-1}}^{h_n} |K(s_n,t)|\,dt > 4^n[2+M(h_{n-1})], \\ &\int_{h_n}^\infty |K(s_n,t)|\,dt < 2^{-n} \end{aligned} \tag{7.66}$$

hold when $n=1$. We continue in this way to obtain sequences $s_1, s_2, \cdots$ and $h_1<h_2<\cdots$, such that $s_n>H$, $h_{n+1}>h_n+1$ and (7.66) holds for each $n=1, 2, 3, \cdots$. Now let $x_0(t)$ be defined by

$$x_0(t) = 2^{-n} \operatorname{sgn} K(s_n,t), \qquad h_{n-1} \leqq t < h_n, \tag{7.67}$$

in which n takes on values $1, 2, \cdots$. Then $x_0 \in K$ by Lemma 4.1; and obviously $|x_0(t)| \leqq 1$ and $x_0(t)\to 0$, so that $x_0 \in K_0$. Hence (7.61) and Theorem 7.3 imply that the transform $y_0(s)$ of $x_0(s)$ exists. The inequalities

$$\left|\int_{h_{n-1}}^{h_n} K(s_n,t)x_0(t)dt\right| = 2^{-n}\int_{h_{n-1}}^{h_n} |K(s_n,t)|\,dt > 2^n[2+M(h_{n-1})],$$

$$\left|\int_0^{h_{n-1}} K(s_n,t)x_0(t)dt\right| \leqq \int_0^{h_{n-1}} |K(s_n,t)|\,dt \leqq M(h_{n-1}),$$

$$\left|\int_{h_n}^\infty K(s_n,t)x_0(t)dt\right| \leqq \int_{h_n}^\infty |K(s_n,t)|\,dt \leqq 2^{-n}$$

imply that

$$| y_0(s_n) | > 2^n[2 + M(h_{n-1})] - M(h_{n-1}) - 2^{-n} > 2^n,$$

and hence

$$\lim_{n\to\infty} | y_0(s_n) | = \infty . \tag{7.68}$$

Thus y_0 does not have a bounded transform. It follows that at least one of the two real functions $\mathcal{R}x_0$ and $\mathfrak{I}x_0$ (which are both members of $\mathcal{R}K_0$ and have transforms) must fail to have a bounded transform and Theorem 7.4 is proved.

The proof of Theorem 7.5 is the same as that of Theorem 7.4 except that Theorem 7.2 is used to obtain (7.61), and that in (7.62), (7.63), and (7.64), "l.u.b.$_{s>H}$" is replaced by "lim sup$_{s\to\infty}$" so that we can choose the sequence $s_1, s_2, \cdots$ such that $s_{n+1} > s_n + 1$ and obtain a contradiction of (7.52).

8. **Application of K to functions $x \in K_B$.** Let K_B denote the class of functions $x(t) \in K$ which are bounded over $0 < t < \infty$. In this section, we give two closely related theorems.

THEOREM 8.1. *If K is such that*

$$\int_0^\infty | K(s, t) | \, dt < \infty , \qquad s > 0, \tag{8.11}$$

and

$$\lim_{h\to\infty} \text{l.u.b.}_{s>0} \int_h^\infty | K(s, t) | \, dt = \Lambda , \tag{8.12}$$

then (whether Λ is finite or $+\infty$) there is a function $x \in K_B$ such that

$$| x(t) | \leqq 1 , \qquad t \geqq 0, \tag{8.13}$$

and the transform $y(s)$ of $x(t)$ exists for $s > 0$ and has the property

$$\text{l.u.b.}_{u,v>0} | y(u) - y(v) | \geqq 2\Lambda . \tag{8.14}$$

Moreover if $K(s, t)$ is real, the function $x(t)$ may be taken real.

THEOREM 8.2. *If K is such that*

$$\int_0^\infty | K(s, t) | \, dt < \infty , \qquad s > 0, \tag{8.21}$$

and

$$\lim_{h\to\infty} \limsup_{s\to\infty} \int_h^\infty |K(s,t)|\,dt = \Lambda, \tag{8.22}$$

then (where Λ is finite or $+\infty$) there is a function $x \in K_B$ such that

$$|x(t)| \leqq 1, \qquad t \geqq 0, \tag{8.23}$$

and the transform $y(s)$ of $x(t)$ exists and has the property

$$\limsup_{u,v\to\infty} |y(u) - y(v)| \geqq 2\Lambda. \tag{8.24}$$

Moreover if $K(s,t)$ is real, the function $x(t)$ may be taken real.

We first prove Theorem 8.1. The case $\Lambda = 0$ is trivial. If $\Lambda = +\infty$, the left member of (7.43), with $H = 0$, will be $+\infty$; hence, by Theorem 7.4, $x \in \mathcal{R}K_0$ exists such that (7.42) fails. But (8.11) implies that the transform $y(s)$ of $x(t)$ exists; hence l.u.b. $|y(s)| = +\infty$ and (8.14) follows. We consider now the remaining case in which $0 < \Lambda < \infty$.

The hypothesis (8.12) implies that we can choose $h_1 > 0$ such that

$$\Lambda - \frac{1}{n} < \text{l.u.b.}_{s>0} \int_{h_n}^\infty |K(s,t)|\,dt < \Lambda + \frac{1}{n} \tag{8.31}$$

will hold when $n = 1$, and then choose $s_1 > 0$ such that

$$\Lambda - \frac{1}{n} < \int_{h_n}^\infty |K(s_n,t)|\,dt < \Lambda + \frac{1}{n} \tag{8.32}$$

holds when $n = 1$. Next we can choose $h_2 > h_1 + 1$ such that

$$\int_{h_n+1}^\infty |K(s_n,t)|\,dt < \frac{1}{n} \tag{8.33}$$

holds when $n = 1$, and (8.31) holds when $n = 2$, and then choose $s_2 > 0$ such that (8.32) holds when $n = 2$. We can continue by induction to define sequences h_n and s_n such that $h_{n+1} > h_n + 1$, $s_n > 0$, and (8.31), (8.32), and (8.33) hold for each $n = 1, 2, \cdots$. From (8.32) and (8.33) we obtain

$$\Lambda - \frac{2}{n} < \int_{h_n}^{h_{n+1}} |K(s_n,t)|\,dt < \Lambda + \frac{2}{n}, \qquad n = 1, 2, \cdots. \tag{8.34}$$

Let $x(t) = 0$ when $0 \leqq t < h_1$; and for each $n = 1, 2, \cdots$ let

$$x(t) = \sigma_n \operatorname{sgn} K(s_n,t), \qquad h_n \leqq t < h_{n+1}, \tag{8.35}$$

where each σ_n is one of the numbers $+1$ or -1 to be determined pres-

ently. Then however we determine the σ_n, $|x(t)| \leqq 1$; $x \in K$ by Lemma 4.1; and our hypothesis and Theorem 7.3 imply that the transform $y(s)$ of $x(t)$ exists. For each $n=1, 2, \cdots$ we find

$$y(s_n) = A_n + \sigma_n \Lambda + R_n, \tag{8.36}$$

where

$$A_n = \int_0^{h_n} K(s_n, t)x(t)dt = \int_{h_1}^{h_n} K(s_n, t)x(t)dt, \tag{8.37}$$

and

$$R_n = \sigma_n\left[\int_{h_n}^{h_{n+1}} |K(s_n, t)|\,dt - \Lambda\right] + \int_{h_{n+1}}^{\infty} K(s_n, t)x(t)dt. \tag{8.38}$$

Using (8.31), we find

$$|A_n| \leqq \int_{h_1}^{\infty} |K(s_n, t)|\,dt < \Lambda + 1, \tag{8.39}$$

and using (8.34) and (8.33) we find

$$\lim_{n\to\infty} R_n = 0. \tag{8.41}$$

If we set $B_n = \mathcal{R}A_n$ and $C_n = \mathcal{R}R_n$, then

$$\mathcal{R}y(s_n) = B_n + \sigma_n\Lambda + C_n. \tag{8.42}$$

If we set

$$X_n = B_n + \sigma_n\Lambda, \tag{8.43}$$

then (8.41), (8.42), (8.43), and the definitions of B_n and C_n imply that

$$\begin{aligned} \text{l.u.b.}_{u,v>0} |y(u) - y(v)| &\geqq \limsup_{m,n\to\infty} |y(s_m) - y(s_n)| \\ &\geqq \limsup_{m,n\to\infty} |\mathcal{R}y(s_m) - \mathcal{R}y(s_n)| \\ &= \limsup_{m,n\to\infty} |X_m - X_n|. \end{aligned} \tag{8.44}$$

Accordingly it is sufficient to determine the $\sigma_n = \pm 1$ so that

$$\limsup_{m,n\to\infty} |X_m - X_n| \geqq 2\Lambda. \tag{8.45}$$

The definition of A_n shows that $A_1 = 0$ and hence $B_1 = \mathcal{R}A_1 = 0$, and that for each $n=1, 2, \cdots$ the constants $\sigma_1, \sigma_2, \cdots, \sigma_n$ determine A_{n+1} and hence $B_{n+1} = \mathcal{R}A_{n+1}$. Moreover (8.39) and the definition

$B_n = \mathcal{R}A_n$ imply that, however the σ_n are defined, the inequalities

$$-(\Lambda+1) < B_n < \Lambda+1, \qquad n = 1, 2, \cdots, \tag{8.46}$$

must hold.

We now indicate, by giving one step, how it is possible to define the σ_n by induction to obtain (8.45). Suppose σ_n has been defined for $n < n_p$ and accordingly B_n is determined for $n \leqq n_p$, p and n_p being positive integers. To simplify typography, let $q = n_p$. Let the interval $-(\Lambda+1) \leqq t < (\Lambda+1)$ in which all points B_n must lie be divided into p equal subintervals $I_1, I_2, \cdots, I_p$, each closed on the left and open on the right. Let $I^{(1)}$ denote the subinterval containing B_q. Let $\sigma_q = 1$, and let $I^{(2)}$ benote the subinterval containing B_{q+1}. If $I^{(2)} \neq I^{(1)}$, let $\sigma_{q+1} = 1$, but if $I^{(2)} = I^{(1)}$, let $\sigma_{q+1} = -1$; and then let $I^{(3)}$ denote the subinterval containing B_{q+2}. If $I^{(3)}$ differs from both $I^{(1)}$ and $I^{(2)}$, let $\sigma_{q+2} = 1$; otherwise let $\sigma_{q+2} = -1$. We continue in this manner to define σ_n for $n_p \leqq n < n_{p+1}$ where $n_{p+1} = n_p + p + 1$. Since the number $(p+1)$ of points B_n is greater than the number p of subintervals, there must be two indices α and β with $p_n \leqq \alpha < \beta < p_{n+1}$ such that

$$|B_\alpha - B_\beta| < 2(\Lambda+1)/p, \tag{8.47}$$

and $\sigma_\alpha = +1$, $\sigma_\beta = -1$, so that

$$X_\alpha = B_\alpha + \Lambda, \qquad X_\beta = B_\beta - \Lambda. \tag{8.48}$$

The two inequalities (8.47) and (8.48) imply

$$X_\alpha - X_\beta > 2\Lambda - 2(\Lambda+1)/p. \tag{8.49}$$

The indices α and β in (8.49) depend upon p, and as p becomes infinite, so also do α and β. Hence (8.49) implies (8.45) and therefore (8.14). Finally if $K(s, t)$ is real, the function $x(t)$ which we constructed is real and Theorem 8.1 is proved.

From (8.44) and (8.45), we obtain

$$\text{l.u.b.}_{u,v>0} |\mathcal{R}y(u) - \mathcal{R}y(v)| \geqq 2\Lambda; \tag{8.51}$$

hence we have proved the stronger theorem obtained by replacing (8.14) by (8.51) in Theorem 8.1. However this extension follows from Theorem 8.1 itself. For if $x(t)$ is a function of the required type whose transform satisfies (8.13), then we can choose a real angle θ such that $x(t)e^{i\theta}$ will be a function of the required type whose transform satisfies (8.51).

We turn now to proof of Theorem 8.2. The case $\Lambda = 0$ is trivial, and the case $\Lambda = \infty$ is covered by Theorem 7.5. In case $0 < \Lambda < \infty$, the hy-

pothesis (8.22) enables us to choose a number $N>0$ such that

$$\int^{\infty} |K(s,t)|\,dt < \Lambda + 1, \qquad h, s \geqq N. \tag{8.52}$$

We can then choose sequences $N<h_1<h_2<\cdots$ and $N<s_1<s_2<\cdots$ such that $s_{n+1}>s_n+1$, $h_{n+1}>h_n+1$, the estimate obtained by replacing "l.u.b." by "lim sup" in (8.31) holds, and (8.32) and (8.33) hold. The proof is then precisely like that of Theorem 8.1 except that (8.39) is implied by (8.52) rather than the analogue of (8.22), and that "l.u.b.$_{u,v>0}$" is replaced by "lim sup$_{u,v\to\infty}$" in (8.44). We thus obtain (8.24) and Theorem 8.2 is proved.

THEOREM 8.6. *If C is a nonnegative constant, then a necessary and sufficient condition that*

$$\text{l.u.b.}_{s>0} |y(s)| \leqq C \text{ l.u.b.}_{s>0} |x(s)|, \qquad x \in K, \tag{8.61}$$

is that

$$\text{l.u.b.}_{s>0} \int_0^{\infty} |K(s,t)|\,dt \leqq C.$$

Necessity is obtained by consideration of the different functions $x(t)=\operatorname{sgn} K(s,t)$ obtained by giving different values to s, and sufficiency is easily established.

9. **Conditions that K be conservative, conservative for null sequences, multiplicative, regular, regular for null sequences, coercive, and null.** A transformation K is called *conservative* (convergence-preserving) *over a class C* of functions if $x \in C$ and existence of $\lim x(t)$ imply existence of $y(s)$ for $s>0$ and of $\lim y(s)$. Equality of $\lim y(s)$ and $\lim x(t)$ is neither required nor prohibited by this definition. If K is conservative over K, then K is called *conservative*. In the next theorem, we use the phrase "conservative over K" rather than "conservative" in order to facilitate statements of closely related theorems.

THEOREM 9.1. *Necessary and sufficient conditions that K be conservative over K are*

$$\int_0^{\infty} |K(s,t)|\,dt < \infty, \qquad s>0, \tag{9.11}$$

$$\limsup_{s\to\infty} \int_0^{\infty} |K(s,t)|\,dt = M < \infty, \tag{9.12}$$

$$\lim_{s\to\infty}\int_0^h K(s,t)x(t)dt = L_h(x), \qquad h>0,\ x\in K, \tag{9.13}$$

$$\lim_{s\to\infty}\int_0^\infty K(s,t)dt = \rho, \tag{9.14}$$

where M and ρ are constants, and $L_h(x)$ is a constant for each choice of $h>0$, $x\in K$.

Necessity of (9.11) and (9.12) follow respectively from Theorems 7.1 and 7.5. To prove necessity of (9.13), let $x\in K$, let $h>0$ and let $x_0(t)=x(t)$ or 0 according as $0\leqq t\leqq h$ or $t>h$. Since $x_0\in K$ and $x_0(t)\to 0$, our hypotheses imply that the transform

$$y_0(s) = \int_0^\infty K(s,t)x_0(t)dt = \int_0^h K(s,t)x(t)dt \tag{9.15}$$

must converge to some limit [which we may denote by $L_h(x)$] as $s\to\infty$. Necessity of (9.14) results from the fact that the transform of the function $x(t)\equiv 1$ must converge as $s\to\infty$.

To prove sufficiency, let $x\in K$ be such that $\lim_{t\to\infty}x(t)=\lambda$ exists. Let $\epsilon>0$. Choose $h>0$ such that

$$|x(t)-\lambda|<\epsilon, \qquad t\geqq h.$$

Then Theorem 7.3 implies that $y(s)$ exists for $s>0$, and we can use the estimate

$$\begin{aligned}|y(u)-y(v)| \leqq &\left|\int_0^h K(u,t)[x(t)-\lambda]dt - \int_0^h K(v,t)[x(t)-\lambda]dt\right| \\ &+\int_h^\infty |K(u,t)||x(t)-\lambda|\,dt \\ &+\int_h^\infty |K(v,t)||x(t)-\lambda|\,dt \\ &+|\lambda|\left|\int_0^\infty K(u,t)dt - \int_0^\infty K(v,t)dt\right|\end{aligned}$$

and our hypotheses to obtain $\limsup_{u,v\to\infty}|y(u)-y(v)|\leqq 2\epsilon M$. Therefore $\lim_{u,v\to\infty}|y(u)-y(v)|=0$, and the Cauchy criterion for convergence implies existence of $\lim_{s\to\infty}y(s)$.

The theorems used in the proof of Theorem 9.1 are sufficiently general to enable us to replace K successively by RK, K_B and RK_B in the proof of Theorem 9.1 to obtain

THEOREM 9.16. *Theorem* 9.1 *remains true when* K_C *is replaced by* RK_C *or by* K_B *or by* RK_B *in its statement.*

THEOREM 9.17. *The conditions* (9.11), (9.12), *and* (9.13) *are necessary and sufficient that* K *be conservative over the class of null functions in* K_C *(that is, functions* $x \in K_C$ *for which* $\lim x(t) = 0$).

Proof of Theorem 9.17 is the same as that of Theorem 9.1 except that necessity of (9.14) need not be proved and that the condition (9.14) is not needed to establish existence of $\lim y(s)$ when $\lim x(t) = \lambda = 0$. Replacing K_C by RK_C or K_0 or RK_0 gives characterizations of transformations conservative over the classes of real null functions in K_C or bounded null functions in K_C or real bounded null functions in K_C.

If K is conservative over C and such that for convergent functions $x \in C$ the value of $\lim y(s)$ is independent of the value of $x(t)$ on each interval $0 \leqq t \leqq h$, then K is called *multiplicative over* C. If K is multiplicative over K_C, then K is *multiplicative.*

THEOREM 9.2. *Necessary and sufficient conditions that* K *be multiplicative over* K_C *are*

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > 0, \tag{9.21}$$

$$\limsup_{s\to\infty} \int_0^\infty |K(s,t)|\,dt = M < \infty, \tag{9.22}$$

$$\lim_{h\to\infty} \int_0^h K(s,t)x(t)dt = 0, \qquad h > 0,\ x \in K_C, \tag{9.23}$$

$$\lim_{s\to\infty} \int_0^\infty K(s,t)dt = \rho. \tag{9.24}$$

Moreover if K *is multiplicative,* $x \in K_C$, *and* $\lim x(t) = \lambda$, *then* x *is summable* K *to* $\rho\lambda$.

The number ρ given by (9.24) is called the *multiplier* of the transformation. Necessity of (9.21), (9.22), and (9.24) follows from Theorem 9.1. To prove necessity of (9.23), let $x \in K_C$ and $h > 0$ be fixed. Let $x_1(t) = x(t)$ or 0 according as $0 \leqq t \leqq h$ or $t > h$, and let $x_2(t) \equiv 0$. Since K is multiplicative, the transforms $y_1(s)$ and $y_2(s)$ of x_1 and x_2 must have the same limit; hence

$$y_1(s) - y_2(s) = \int_0^h K(s,t)x(t)dt$$

must converge to 0 as $s \to \infty$ and proof of necessity is complete. To prove sufficiency, let $x(t) \in K$ and let $\lim_{t\to\infty} x(t) = \lambda$. Let $\epsilon > 0$. If $h > 0$ is such that $|x(t) - \lambda| < \epsilon$ for $t \geqq h$, then the equality

$$y(s) - \rho\lambda = A(s) + B(s) + \int_h^\infty K(s, t)[x(t) - \lambda]dt,$$

where

$$A(s) = \lambda\left[\int_0^\infty K(s, t)dt - \rho\right], \qquad B(s) = \int_0^h K(s, t)[x(t) - \lambda]dt,$$

implies

$$|y(s) - \rho\lambda| \leqq |A(s)| + |B(s)| + \epsilon\int_h^\infty |K(s, t)|\,dt,$$

and we can use our hypotheses to obtain $\limsup_{s\to\infty} |y(s) - \rho\lambda| \leqq \epsilon M$. Therefore $\lim_{s\to\infty} y(s) = \rho\lambda$, that is, $x(t)$ is summable to $\rho\lambda$. Since $\rho\lambda$ is independent of the values of $x(t)$ over each finite interval $(0, h)$, the theorem is proved. Replacing K by RK, K_B, RK_B in this proof gives

THEOREM 9.25. *Theorem* 9.2 *remains true when* K *is replaced by* RK *or by* K_B *or by* RK_B *in its statement.*

We now prove Theorem 5.1 and related theorems. Let K be regular over C, where C is one of the classes K or RK or K_B or RK_B. Then the definitions of regular and multiplicative transformations imply that K is multiplicative over C. Hence by Theorems 9.2 and 9.25 the function $x_0(t) \equiv 1$ is summable K to ρ. Regularity of K over C implies that $\rho = \lim x_0(t) = 1$. On the other hand Theorems 9.2 and 9.25 imply that if K is multplicative over C and $\rho = 1$, then K is regular over C. Thus the class of transformations regular over C is identical with the class of transformations multiplicative over C with multiplier $\rho = 1$. This fact and Theorems 9.2 and 9.25 prove Theorem 5.1 and the following one.

THEOREM 9.3. *Theorem* 5.1 *remains true when* K *is replaced by* RK (the class of real functions in K) *or by* K_B (the class of bounded functions in K) *or by* RK_B (the class of real bounded functions in K) *in its statement.*

Methods of proof already used in this section suffice to prove the following two theorems.

THEOREM 9.4. *In order that* K *be regular over the class of null functions in* K, *it is necessary and sufficient that*

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > 0, \tag{9.41}$$

$$\limsup_{s\to\infty} \int_0^\infty |K(s,t)|\,dt < \infty, \tag{9.42}$$

$$\lim_{s\to\infty} \int_0^h K(s,t)x(t)dt = 0, \qquad h > 0,\ x \in \mathcal{K}. \tag{9.43}$$

THEOREM 9.44. *Theorem* 9.4 *remains true when* $\mathcal{K}$ *is replaced by* $\mathcal{RK}$ *or* $\mathcal{K}_B$ *or* $\mathcal{RK}_B$ *in its statement.*

A transformation K is called *coercive* if each $x \in \mathcal{K}_B$ (that is, each bounded function $x \in \mathcal{K}$) is summable K.

THEOREM 9.5. *Necessary and sufficient conditions that* K *be coercive are*

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > 0, \tag{9.51}$$

$$\lim_{s\to\infty} \int_0^h K(s,t)x(t)dt = L_h(x), \qquad h > 0,\ x \in \mathcal{K}_B, \tag{9.52}$$

$$\lim_{h\to\infty} \limsup_{s\to\infty} \int_h^\infty |K(s,t)|\,dt = 0. \tag{9.53}$$

Necessity of (9.51) and (9.52) follows from Theorem 9.15 and the fact that each coercive transformation is conservative over $\mathcal{K}_B$. Necessity of (9.53) follows from Theorem 8.2; for, if the left member of (9.53) is $\Lambda > 0$, then there is a function $x \in \mathcal{K}_B$ for which (8.24) holds, and therefore x is not summable K. To establish sufficiency, let $x \in \mathcal{K}$ be such that l.u.b. $|x(t)| < \lambda < \infty$. Let $\epsilon > 0$. Choose $h > 0$ and $s_0 > 0$ such that

$$\int_h^\infty |K(s,t)|\,dt < \epsilon/2\lambda, \qquad s \geqq s_0.$$

The hypothesis (9.51) guarantees that the transform $y(s)$ of $x(t)$ exists. Hence

$$|y(u) - y(v)| \leqq |F(u) - F(v)| + |G(u)| + |G(v)|, \qquad u, v > 0,$$

where

$$F(s) = \int_0^h K(s,t)x(t)dt, \qquad G(s) = \int_h^\infty K(s,t)x(t)dt.$$

Our hypotheses and inequalities imply that $|F(u) - F(v)| \to 0$ as u,

$v \to \infty$, and that $G(s) < \epsilon/2$ when $s \geqq s_0$. Hence $\limsup |y(u) - y(v)| \leqq \epsilon$ and existence of $\lim y(s)$ follows.

If K is coercive and such that for each $x \in K_B$ the value of $\lim y(s)$ is independent of the value of $x(t)$ on each finite interval $0 \leqq t \leqq h$, then k is called *null*.

THEOREM 9.6. *Necessary and sufficient conditions that K be null are*

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > 0, \tag{9.61}$$

$$\lim_{s\to\infty} \int_0^h K(s,t)x(t)dt = 0, \qquad h > 0,\ x \in K_B, \tag{9.62}$$

$$\lim_{h\to\infty} \limsup_{s\to\infty} \int_h^\infty |K(s,t)|\,dt = 0. \tag{9.63}$$

Moreover if K is null, then each $x \in K_B$ is summable K to 0.

Necessity follows from the fact that each null transformation is both coercive and multiplicative. Sufficiency is easily established by proving that each $x \in K_B$ is summable K to 0.

THEOREM 9.7. *Let C be a nonnegative real constant. In order that K may be such that the transform $y(s)$ of each $x \in K$ for which* $\limsup |x(t)| < \infty$ *exists and has the property*

$$\limsup_{s\to\infty} |y(s)| \leqq C \limsup_{t\to\infty} |x(t)|, \tag{9.71}$$

it is necessary and sufficient that

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > 0, \tag{9.72}$$

$$\lim_{s\to\infty} \int_0^h K(s,t)x(t)dt = 0, \qquad h > 0,\ x \in K, \tag{9.73}$$

and

$$\lim_{h\to\infty} \limsup_{s\to\infty} \int_h^\infty |K(s,t)|\,dt \leqq C. \tag{9.74}$$

For proof of necessity, the hypotheses imply that K is regular over the class of null functions $x \in K$ and hence that (9.72) and (9.73) hold. Necessity of (9.74) is implied by Theorem 8.2. Sufficiency is easily established. Theorem 9.7 remains true when K is replaced by K_B.

Each theorem of this section has contained the condition

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > 0, \tag{9.81}$$

explicitly, and the condition

$$\limsup_{s\to\infty} \int_0^\infty |K(s,t)|\,dt < \infty \tag{9.82}$$

either explicitly or implicitly. Each of these theorems remains true when (9.81) is deleted from the set of conditions and the definitions of regularity and so on are modified to correspond to the modified definition of summability under which $x(t)$ is called summable to L if

$$y(s) = \int_0^\infty K(s,t)x(t)dt \tag{9.83}$$

exists for all sufficiently great s [that is, for all $s \geqq s_0$ where s_0 may depend upon the particular function x in (9.83)] and $\lim_{s\to\infty} y(s) = L$.* Proof of necessity for the new theorems is identical with that for the old, the theorems of §§7 and 8 having been made sufficiently general to make this true. Condition (9.82) implies existence of H such that

$$\int_0^\infty |K(s,t)|\,dt < \infty, \qquad s > H, \tag{9.84}$$

and this plays the role of (9.81) in proof of sufficiency for the new theorems.

10. **Kernels which are step functions of t for each s; sequence to function transformations; matrix transformations.** It was pointed out in the introduction that a matrix transformation A can be identified with a kernel transformation K whose domain and range are confined to the class of (or a subclass of) step functions which are constant over each unit interval $n \leqq t < n+1$, $n = 0, 1, 2, \cdots$. This fact is, of course, significant, but it does not imply that from each theorem involving matrix transformations follows ipso facto a corresponding theorem involving kernel transformations whose domain and range are not restricted to classes of step functions. For example, matrix transformations include the identity transformation

$$y_s = \sum_{t=0}^{\infty} \delta_{s,t} x_t, \qquad s = 0, 1, 2, \cdots,$$

* This modified definition of "summable" is used in a study of regular transformations by J. D. Tamarkin, *On the notion of regularity of methods of summation of infinite series*, this Bulletin, vol. 41 (1935), pp. 241–243.

in which $\delta_{s,t}$ is 0 or 1 according as $t \neq s$ or $t = s$; but there is no kernel transformation

$$y(s) = \int_0^\infty K(s, t)x(t)dt$$

such that $y(s) = x(s)$, $0 < s < \infty$, for all functions $x(s)$. It is likewise true that theorems involving kernel transformations do not ipso facto imply corresponding theorems involving matrix transformations.

The theorems of §9 involving kernel transformations are not merely analogous to, but actually imply, corresponding results for sequence-to-function transformations and for sequence-to-sequence matrix transformations.

A sequence $A_0(s), A_1(s), \cdots$ of complex-valued functions defined for $s > 0$ determines the *sequence-to-function transformation*

$$y(s) = \sum_{t=0}^{\infty} A_t(s)x_t, \qquad s > 0, \tag{10.01}$$

which associates with each sequence $x_0, x_1, \cdots$ for which the series converges for $s > 0$ a transform $y(s)$. The transformation (10.01) is *regular* if $\lim x_t = L$ implies $\lim y(s) = L$, and other definitions are analogous to those for matrix transformations. Let x_t be any sequence of complex numbers. Then, assuming that $\int$ has the properties of §2,

$$\sum_{t=0}^{n} A_t(s)x_t = \sum_{k=0}^{n-1} \int_k^{k+1} A_k(s)x_k dt, \qquad n = 0, 1, 2, \cdots. \tag{10.02}$$

If we define $K(s, t)$ and $x(t)$ by the formulas (in which k takes values $0, 1, 2, \cdots$)

$$K(s, t) = A_k(s), \qquad k \leqq t < k + 1, \ s > 0, \tag{10.03}$$

$$x(t) = x_k, \qquad k \leqq t < k + 1, \tag{10.04}$$

then K is a kernel according to the definition of §3, and $x \in \mathfrak{K}$. The right member of (10.02) can then be written in the forms

$$\sum_{k=0}^{n-1} \int_k^{k+1} K(s, t)x(t)dt = \int_0^n K(s, t)x(t)dt. \tag{10.05}$$

Setting, for all integers $n \geqq 0$, all real $h \geqq 0$, and all $s > 0$,

$$F(n, s) = \sum_{t=0}^{n} A_t(s)x_t, \qquad G(h, s) = \int_0^h K(s, t)x(t)dt,$$

we see that, for each s, $G(h, s)$ is a linear function of h over each closed

interval $n \leqq h \leqq n+1$. Moreover, since $F(n, s) = G(n, s)$, it follows that existence of either one of

$$(10.06) \qquad y^{(A)}(s) = \sum_{t=0}^{\infty} A_t(s)x_t = \lim_{n\to\infty} \sum_{t=0}^{n} A_t(s)x_t$$

or

$$(10.07) \qquad y^{(K)}(s) = \int_0^{\infty} K(s, t)x(t)dt = \lim_{h\to\infty} \int_0^{h} K(s, t)x(t)dt$$

implies existence of the other and the equality $y^{(A)}(s) = y^{(K)}(s)$.

Leaving consideration of other properties to the reader, we discuss regularity. It is clear from the above discussion of (10.06) and (10.07) that, if K is regular, and hence also regular over the class $\mathfrak{K}'$ of step functions $x(t)$ constant over each interval $n \leqq t < n+1$, then (10.07) is regular; and it is also clear that if (10.07) is regular, then K is regular over $\mathfrak{K}'$. Hence the first of the two next theorems is implied by the second.

THEOREM 10.1. *The sequence-to-function transformation*

$$(10.11) \qquad y(s) = \sum_{t=0}^{\infty} A_t(s)x_t, \qquad s > 0,$$

is regular if and only if the step kernel defined by (10.03) *is regular.*

THEOREM 10.2. *If K is a step kernel, $K(s, t)$ being independent of t over each interval $n \leqq t < n+1$, and if K is regular over the class $\mathfrak{K}'$ of step functions constant over each interval $n \leqq t < n+1$, then K is regular.*

To show that K is regular, let $x_0 \in \mathfrak{K}$ and $\lim x_0(t) = L$. We are to show that

$$(10.21) \qquad y_0(s) = \int_0^{\infty} K(s, t)x_0(t)dt, \qquad s > 0,$$

exists and $\lim y_0(s) = L$. Our hypothesis implies, by Theorem 7.15, that

$$(10.22) \qquad \int_0^{\infty} |K(s, t)|\, dt < \infty, \qquad s > 0,$$

and existence of (10.21) follows. Let $A_n(s)$ denote the value of $K(s, t)$ in the interval $n \leqq t < n+1$. The hypothesis $x_0 \in \mathfrak{K}$ implies that $x(t)$ is integrable over each interval $n \leqq t \leqq n+1$ corresponding to an integer n such that $A_n(s) \neq 0$ for some $s > 0$; the condition $x_0 \in \mathfrak{K}$ implies no

condition whatever on $x_0(t)$ in an interval in which $A_n(s)=0$ for all $s>0$, since, in this case,

$$\int_n^{n+1} K(s,t)x_0(t)dt = \int_n^{n+1} A_n(s)x_0(t)dt = 0, \tag{10.23}$$

however $x_0(t)$ is defined. Let $x_1(t)=x_0(t)$ in each interval $n\leqq t<n+1$ over which $x_0(t)$ is integrable, and let $x_1(t)=L$ for all other $t\geqq 0$. Then $x_1 \varepsilon K$ and $\lim x_1(t)=L$. Let

$$x_2(t) = \int_n^{n+1} x_1(u)du, \qquad n \leqq t < n+1,\ n = 0, 1, \cdots. \tag{10.24}$$

Then $x_2 \varepsilon K'$ and $x_2 \to L$. Starting with (10.21), we can show that

$$\begin{aligned} y_0(s) &= \lim_{m\to\infty} \sum_{n=0}^{m-1} \int_n^{n+1} K(s,t)x_0(t)dt = \lim_{m\to\infty} \sum_{n=0}^{m-1} \int_n^{n+1} A_n(s)x_0(t)dt \\ &= \lim_{m\to\infty} \sum_{n=0}^{m-1} \int_n^{n+1} A_n(s)x_1(t)dt = \lim_{m\to\infty} \sum_{n=0}^{m-1} A_n(s)x_2(n) \\ &= \lim_{m\to\infty} \sum_{n=0}^{m-1} \int_n^{n+1} A_n(s)x_2(t)dt = \int_0^{\infty} K(s,t)x_2(t)dt. \end{aligned} \tag{10.25}$$

But since $x_2 \varepsilon K'$ and K is regular over K', it follows that the last member of (10.25), and hence $y_0(s)$, converges to L as $s\to\infty$. This proves Theorem 10.2 and hence also Theorem 10.1.

THEOREM 10.3. *Necessary and sufficient conditions that the sequence-to-function transformation*

$$y(s) = \sum_{t=0}^{\infty} A_t(s)x_t \tag{10.31}$$

be regular are

$$\sum_{t=0}^{\infty} |A_t(s)| < \infty, \qquad s > 0, \tag{10.32}$$

$$\limsup_{s\to\infty} \sum_{t=0}^{\infty} |A_t(s)| = M < \infty, \qquad s > 0, \tag{10.33}$$

$$\lim_{s\to\infty} A_t(s) = 0, \qquad s = 0, 1, 2, \cdots, \tag{10.34}$$

$$\lim_{s\to\infty} \sum_{t=0}^{\infty} A_t(s) = 1. \tag{10.35}$$

When K is the step kernel given by (10.03), the four criteria (5.11), (5.12), (5.13), and (5.14) for regularity of K are easily shown to be equivalent respectively to (10.32), (10.33), (10.34), and (10.35). Hence Theorem 10.3 is implied by Theorems 5.1 and 10.1.

It is often stated that the set of conditions consisting of

$$\sum_{t=0}^{\infty} |A_t(s)| < M < \infty, \qquad s > 0, \tag{10.36}$$

(10.34) and (10.35) is necessary and sufficient for regularity of (10.31); this statement is false since the transformation

$$\begin{aligned} y(s) &= x_0, && s = 0, \\ &= (1/s)x_0 + x_{[s]}, && s > 0, \end{aligned}$$

is regular and fails to satisfy (10.36). On one hand, (10.36), (10.34), and (10.35) are sufficient for regularity, since (10.36) implies (10.32) and (10.33). On the other hand, the error in assuming that they are also necessary for regularity does not usually lead to false results on account of the fact that if

$$y(s) = \sum_{t=0}^{\infty} A_t(s)x_t, \qquad s > 0,$$

is a regular transformation which fails to satisfy (10.36), then there is a constant $H>0$ such that the transformation

$$y(s) = \sum_{t=0}^{\infty} A_t(s)x_t, \qquad s > H,$$

satisfies the condition

$$\sum_{t=0}^{\infty} |A_t(s)| < M < \infty, \qquad s > H,$$

which is analogous to (10.36).

If $A_{s,t}$ is a matrix of complex constants, and we define functions $A_t(s)$ by the formulas

$$A_t(s) = A_{n,t}, \qquad n \leqq s < n + 1; n = 0, 1, \cdots,$$

it becomes apparent that the matrix transformation

$$y_s = \sum_{t=0}^{\infty} A_{s,t}x_t \tag{10.37}$$

is regular if and only if (10.31) is regular. Hence, using Theorem 10.3,

we see that necessary and sufficient conditions for regularity of matrix transformations are those stated at the beginning of §5.

Conditions that (10.31) and (10.37) be regular over other classes of sequences (for example, null sequences) are obtained by slight alteration of a few parts of this section.

11. **Scope of regular transformations.** Steinhaus* proved that no regular matrix A exists which evaluates all bounded sequences x_t. The following theorem extends this result to kernel transformations.

THEOREM 11.1. *If K is a regular transformation, then there is a function $x \in K$ for which $|x(t)| \leqq 1$ while the transform $y(s)$ has the property*

$$\limsup_{u,v\to\infty} |y(u) - y(v)| \geqq 2. \tag{11.11}$$

This theorem is an immediate result of Theorems 5.2 and 8.2.

THEOREM 11.2. *If, under a given definition of integral having the properties listed in §2, $x(t)$ is a bounded function which is integrable over each finite interval $0 \leqq t \leqq h$, then there is a regular transformation K involving the given definition of integral such that $x(t)$ is summable K.*

The manner in which we construct a regular transformation which evaluates $x(t)$ indicates that many such transformations can be constructed. The hypotheses of Theorem 10.2 imply that the sequence $z(0), z(1), \cdots$ of complex numbers defined by

$$z(n) = \int_n^{n+1} x(t)dt, \qquad n = 0, 1, 2, \cdots,$$

is bounded. Hence there is a sequence $0 = n_1 < n_2 < \cdots$ of indices such that $\lim_{p\to\infty} z(n_p)$ exists. The transformation

$$y(s) = \int_{n_p}^{n_{p+1}} x(t)dt, \qquad n_p \leqq s < n_{p+1},$$

is obviously regular and evaluates $x(t)$ to $\lim z(n_p)$.

Theorems 11.1 and 11.2 indicate that the fitting of regular transformations to functions is like the fitting of shoes to men. No one pair of shoes will fit all men; but any man can obtain from a shoemaker many different styles of shoes which will fit him.

12. **Conclusion.** Let E and $\mathcal{E}$ be point sets, and let integration be so defined that for at least some complex-valued functions $f(t)$ defined

* H. Steinhaus, *Some remarks on the generalizations of the notion of limit* (in Polish), Prace Matematyczno-Fizyczne, vol. 22 (1921), pp. 121–134.

for $t \in E$, the symbol

$$\int_E f(t)dt \tag{12.1}$$

represents a complex number corresponding to f. There is now the possibility that if the integral has appropriate properties, and if $K(s, t)$ is an appropriate complex-valued function defined for $s \in \mathcal{E}$, $t \in E$, then the transformation

$$y(s) = \int_E K(s, t)x(t)dt, \qquad s \in \mathcal{E}, \tag{12.2}$$

is significant. In case the set $\mathcal{E}$ is one (not necessarily in any euclidean space or even in a metric space) in which the notion of limit point is defined, and s_0 is a limit point of $\mathcal{E}$, then $x(t)$ may be called *summable* (12.2) to L if $y(s)$ exists for $s \in \mathcal{E}$ and $y(s) \to L$ as $s \to s_0$ over $\mathcal{E}$. In case t_0 is a limit point of E, the number $\lim_{s \to s_0} y(s)$ may (when it exists) be regarded as a *generalized limit* of $x(t)$ as $t \to t_0$ over E. The transformation (12.2) is *regular over a class* C of functions if $x(t)$ is summable to L whenever $x \in C$ and $x(t) \to L$ as $t \to t_0$ over E. Other properties which (12.2) may have or fail to have can be similarly defined.

The transformation (12.2) and the associated method of summability determined by t_0 and s_0 reduce to the form we have studied when $\mathcal{E}$ is the set $0 < s < \infty$, E is the set $0 < t < \infty$, and t_0 and s_0 are the symbolic limit points $+\infty$.

It is apparent that the role of t in the theory of the transformation (12.2) lies far deeper than the role of s. In so far as the theory of K which we have covered in this address is concerned, the difference between K and the transformation

$$y(s) = \int_0^\infty K(s, t)x(t)dt, \qquad s \in \mathcal{E}, \tag{12.3}$$

with $s \to s_0$ instead of $s \to \infty$, is trivial. For example, (12.3) may furnish a regular generalization of $\lim_{t \to \infty} x(t)$ when $s \to s_0$ over a set $\mathcal{E}$ which is a bounded interval in euclidean space of one dimension, a "curve" in a plane, or a non-measurable set in 3-space.

On the other hand the transformation

$$y(s) = \int_E K(s, t)x(t)dt, \qquad s > 0, \tag{12.4}$$

in which $s_0 = +\infty$, assumes entirely different forms according to the

character of E and t_0 and the definition of integral used. In case E is the interval $a<t<\infty$, where a is a constant not 0, the transformation (12.4) differs only in an obvious and trivial way from K. Easy modifications of our theory of K cover some cases in which E is a curve other than the interval $a<t<\infty$ in a plane. These examples furnish generalizations of $\lim x(t)$ as $t\to t_0$ over a curve.

For an example of a different character, let E denote the set of points of a plane which are interior to the unit circle with center at the origin, and let t_0 be a point on the unit circle. For each $h>0$, let E_h be the set of points interior to the circle with center at t_0 and radius h. For each $h>0$, let

$$\int_{E-E_h} f(t)dt$$

denote the Lebesgue double integral of f over $(E-E_h)$ when the integral exists, and let

$$\int_E f(t)dt = \lim_{h\to 0} \int_{E-E_h} f(t)dt$$

when the limit exists. The transformation (12.4) then furnishes a generalized definition of the limit of $x(t)$ as $t\to t_0$ over the interior of the unit circle.

It may be possible to give a theory of (12.2) covering at the same time all of the special cases which we have mentioned and many others. Such a theory would be more impressive and would have wider application than the theory we have given. However the details necessary to make the more general theory precise would have made our discussion more cumbersome, and might have obscured the real purpose of this address.

CORNELL UNIVERSITY

NEW POINT CONFIGURATIONS AND ALGEBRAIC CURVES CONNECTED WITH THEM*

ARNOLD EMCH

1. **Introduction.** In the memorial volume† for Professor Hayashi, I studied an involutorial Cremona transformation in a projective S_r which is obtained as follows: Let $C_i=(ax)_i\lambda_i^2+(bx)_i\lambda_i+(cx)_i=0$, $(i=1, 2, \cdots, r)$, be r hypercones in S_r. Every value of λ_i determines a hypertangent plane to the cone C_i. Thus the parameters $\lambda_1, \lambda_2, \cdots, \lambda_r$ for the hypercones $C_1, C_2, \cdots, C_r$, in the same order, determine r hyperplanes which intersect in a point (x) of S_r. From this point (x) there pass, one for each of the r hypercones, r more tangent hyperplanes whose parameters $\lambda_1', \lambda_2', \cdots, \lambda_r'$ are in the same order uniquely determined by the set $\lambda_1, \lambda_2, \cdots, \lambda_r$, and hence are rational functions

$$\rho\lambda_i' = \phi_i(\lambda_1, \lambda_2, \cdots, \lambda_r), \qquad i = 1, 2, \cdots, r,$$

of the parameters λ. Conversely, the set $\lambda_1', \lambda_2', \cdots, \lambda_r'$ determines λ_i uniquely: $\sigma\lambda_i=\phi_i(\lambda_1', \lambda_2', \cdots, \lambda_r')$. If therefore the λ's and λ''s are interpreted as coordinates of points of euclidean spaces $E_r(\lambda)$ and $E_r'(\lambda')$, there exists an involutorial Cremona transformation between the two r-dimensional spaces. The order and fundamental elements of this involution were determined in the corresponding projective spaces S_r and S_r' and applications given for S_2 and S_3. These belong to a remarkable class of involutions which have the property that when in S_r and S_r'

$$P(\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_{r+1}), \qquad P'(\lambda_1', \lambda_2', \lambda_3', \cdots, \lambda_{r+1}')$$

are corresponding points and any number of transpositions between coordinates in the same columns is performed, say

$$Q(\lambda_1, \lambda_2', \lambda_3', \cdots, \lambda_i', \cdots, \lambda_r, \cdots, \lambda_{r+1}'),$$

$$Q'(\lambda_1', \lambda_2, \lambda_3, \cdots, \lambda_i, \cdots, \lambda_r', \cdots, \lambda_{r+1}),$$

then Q, Q' is always a couple of corresponding points of the involution.

To this class also belong the well known quadratic and cubic involutions in $S_2, \rho x_i' = 1/x_i$, $(i=1, 2, 3)$, and in $S_3, \rho x_i' = 1/x_i$, $(i=1, 2, 3, 4)$,

* Presented to the Society, September 6, 1938.

† The Tôhoku Mathematical Journal, vol. 37 (1933), pp. 100–109. See also Commentarii Mathematici Helvetici, vol. 4 (1932), pp. 65–73.

and in general in S_r, $\rho x_i' = 1/x_i$, $(i=1, 2, \cdots, r+1)$. It is the purpose of this paper to show the importance of these in connection with the plane elliptic cubic in S_2 and a certain septimic of genus three in S_3.

2. **The Δ_8-configuration on the plane elliptic cubic.** Let A_1 (1, 0, 0); A_2 (0, 1, 0); A_3 (0, 0, 1); B (1, 1, 1); B_1 (−1, 1, 1); B_2 (1, −1, 1); B_3 (1, 1, −1) be the fundamental and invariant points of the quadratic involution in S_2, $T_2 \equiv \rho x_i' = 1/x_i$, $(i=1, 2, 3)$, and perform the possible permutations between the coordinates of corresponding points as indicated above, so that we obtain the four couples of corresponding points

$$P_1\ (a_1, a_2, a_3), \qquad P_2\ (a_2a_3, a_2, a_3),$$
$$P_1'\ (a_2a_3, a_3a_1, a_1a_2), \qquad P_2'\ (a_1, a_3a_1, a_1a_2),$$
$$P_3\ (a_1, a_3a_1, a_3), \qquad P_4\ (a_1, a_2, a_1a_2),$$
$$P_3'\ (a_2a_3, a_2, a_1a_2), \qquad P_4'\ (a_2a_3, a_3a_1, a_3).$$

This is easily verified. Take for example P_2; its inverse is $(1/a_2a_3, 1/a_2, 1/a_3)$. Multiplying by $a_1a_2a_2$, we obtain as the same point (a_1, a_3a_1, a_1a_2) which is P_2'. The eight points of this Δ_8 configuration can be grouped into couples whose joins pass through A_1, A_2, A_3 as indicated in the following table:

$$A_1\colon\ P_1P_2,\ P_1'P_2',\ P_3P_4',\ P_3'P_4;$$
$$A_2\colon\ P_1P_3,\ P_1'P_3',\ P_2P_4',\ P_2'P_4;$$
$$A_3\colon\ P_1P_4,\ P_1'P_4',\ P_2P_3',\ P_2'P_3.$$

It is moreover evident that the joins of corresponding points P_iP_i' pass through the point O $(a_1+a_2a_3, a_2+a_3a_1, a_3+a_1a_2)$. The result may be stated as the following theorem:

THEOREM 1. *The eight points of a Δ_8-configuration lie by twos on four lines through each of the points A_i. The joins of the four pairs of corresponding points pass through a fixed point O, uniquely determined by one pair of corresponding points in a Δ_8.*

Now O is the isologue of an invariant elliptic cubic C_3 of the involution, uniquely determined by any pair P_iP_i' of Δ_8. Of course, C_3 passes through Δ_8, B, B_1, B_2, B_3, and also O' (a_2a_3, a_3a_1, a_1a_2). Draw any other secant s through O cutting C_3 in a pair Q_1Q_1' of corresponding points. These determine a new Δ_8-configuration on the same C_3, so that there are ∞^1 Δ_8-configurations on C_3 attached to the triangle $A_1A_2A_3$.

On any elliptic cubic C_3 there are ∞^1 Steinerian quadruples (points of tangency of four tangents from a point of C_3 to C_3) and for each a diagonal triangle $A_1A_2A_3$. Choose $A_1A_2A_3$ as the fundamental triangle* of a quadratic transformation and the Steinerian quadruple as the set of invariant points B, B_1, B_2, B_3. The elliptic cubic C_3 is invariant in this involution and the tangents at the B's cut C_3 in its isologue O. Hence we have the following theorem:

THEOREM 2. *There are ∞^1 Δ_8-configurations on each of the diagonal triangles of the ∞^1 Steinerian quadruples of a given elliptic C_3, completely inscribed in this cubic.*

3. **The Δ_{16}-configuration and a septimic of genus three in S_3 connected with it.** (1) Consider the involutorial cubic transformation T_3: $\rho x_i' = 1/x_i$, $(i=1, 2, 3, 4)$, in an S_3 with the fundamental points A_1 (1, 0, 0, 0), A_2 (0, 1, 0, 0), A_3 (0, 0, 1, 0), A_4 (0, 0, 0, 1), and the invariant points B_i $(\pm 1, \pm 1, \pm 1, \pm 1)$, $(i=1, 2, \cdots, 8)$; and perform all possible permutations, or series of transpositions as explained in §1. In this manner a configuration Δ_{16} of eight couples of corresponding points P_iP_i' is obtained as shown in the table

P_1 (a_1, a_2, a_3, a_4);	P_5 $(a_1, a_2, a_3, a_1a_2a_3)$;
P_1' $(a_2a_3a_4, a_1a_3a_4, a_1a_2a_4, a_1a_2a_3)$;	P_5' $(a_2a_3a_4, a_1a_3a_4, a_1a_2a_4, a_4)$;
P_2 $(a_2a_3a_4, a_2, a_3, a_4)$;	P_6 $(a_2a_3a_4, a_1a_3a_4, a_2, a_3)$;
P_2' $(a_1, a_1a_3a_4, a_1a_2a_4, a_1a_2a_3)$;	P_6' $(a_1, a_2, a_1a_2a_4, a_1a_2a_3)$;
P_3 $(a_1, a_1a_3a_4, a_3, a_4)$;	P_7 $(a_2a_3a_4, a_2, a_1a_2a_4, a_4)$;
P_3' $(a_2a_3a_4, a_2, a_1a_2a_4, a_1a_2a_3)$;	P_7' $(a_1, a_1a_3a_4, a_3, a_1a_2a_3)$;
P_4 $(a_1, a_2, a_1a_3a_4, a_4)$;	P_8 $(a_2a_3a_4, a_2, a_3, a_1a_2a_3)$;
P_4' $(a_2a_3a_4, a_1a_3a_4, a_3, a_1a_2a_3)$;	P_8' $(a_1, a_1a_3a_4, a_1a_2a_4, a_4)$.

As in case of S_2 it may be verified at once that the points of each pair P_iP_i' correspond, and that the sixteen points lie by twos on eight lines through each A_i. If Q is any point of S_3, then the line A_iQ is transformed into the line A_iQ'. The lines P_1P_2, P_3P_6 on A_1 and P_1P_3, P_2P_6 on A_2 form a quadrilateral in the plane $a_4x_3 - a_3x_4 = 0$; $P_1'P_2'$, $P_3'P_6'$ on A_1, and $P_1'P_3'$, $P_2'P_6'$ on A_2 a quadrilateral in the conjugate plane $a_3x_3 - a_4x_4 = 0$. The table of the thirty-two lines, eight through each A_i follows:

* For T_2, T_3 and other involutions see the author's paper *On surfaces and curves which are invariant under involutorial Cremona transformations*, American Journal of Mathematics, vol. 48 (1926), pp. 21–44.

$$A_1\colon P_1P_2,\ P_1'P_2',\ P_3P_6,\ P_3'P_6',\ P_4P_7,\ P_4'P_7',\ P_5P_8,\ P_5'P_8';$$
$$A_2\colon P_1P_3,\ P_1'P_3',\ P_2P_6,\ P_2'P_6',\ P_4P_8',\ P_4'P_8,\ P_5P_7',\ P_5'P_7;$$
$$A_3\colon P_1P_4,\ P_1'P_4',\ P_2P_7,\ P_2'P_7',\ P_3P_8',\ P_3'P_8,\ P_5P_6',\ P_5'P_6;$$
$$A_4\colon P_1P_5,\ P_1'P_5',\ P_2P_8,\ P_2'P_8',\ P_3P_7',\ P_3'P_7,\ P_4P_6',\ P_4'P_6.$$

It also appears at once that the eight joins of corresponding points P_iP_i' pass through a fixed point

$$O\,(a_1 + a_2a_3a_4,\ a_2 + a_1a_3a_4,\ a_3 + a_1a_2a_4,\ a_4 + a_1a_2a_3).$$

To sum up we have the following theorem:

THEOREM 3. *The sixteen points of Δ_{16} lie by twos on eight lines through each of the four A_i. The eight lines on each of any two of the four A_i form four quadrilaterals on the chosen two A_i, which lie in two pairs of conjugate planes with the join of the two A_i as a common axis. The joins of corresponding points of Δ_{16} pass through a fixed point O.*

(2) It is known that the system of lines joining corresponding points of T_3 form a cubic line complex Γ:

$$p_{12}p_{13}p_{23} + p_{12}p_{14}p_{42} + p_{13}p_{14}p_{34} + p_{23}p_{42}p_{34} = 0$$

so that the lines of Γ on a point $O\,(b_1, b_2, b_3, b_4)$, $b_1 = a_1 + a_2a_3a_4, \cdots$, generate the cubic complex-cone

$$\begin{aligned} K = {} & (b_1x_2 - b_2x_1)(b_1x_3 - b_3x_1)(b_2x_3 - b_3x_2) \\ & + (b_1x_2 - b_2x_1)(b_1x_4 - b_4x_1)(b_4x_2 - b_2x_4) \\ & + (b_1x_3 - b_3x_1)(b_1x_4 - b_4x_1)(b_3x_4 - b_4x_3) \\ & + (b_2x_3 - b_3x_2)(b_4x_2 - b_2x_4)(b_3x_4 - b_4x_3) = 0. \end{aligned}$$

Among the generatrices of K are the eight joins P_iP_i' of Δ_{16}. The eight lines OB_i lie on K. Any other generatrix g of K is on two corresponding points Q and Q' of T_3. These determine another Δ_{16} uniquely, which also lies on K. Thus there are ∞^1 Δ_{16}'s on K. Corresponding points QQ' on K form a certain space curve whose order is obtained as follows: The join of $Q(x)$, $Q'(x')$ passes through O when

$$\lambda x_1 + \mu x_2x_3x_4 = b_1, \qquad \lambda x_2 + \mu x_1x_3x_4 = b_2,$$
$$\lambda x_3 + \mu x_1x_2x_4 = b_3. \qquad \lambda x_4 + \mu x_1x_2x_3 = b_4.$$

Eliminating λ, μ, 1 from any two distinct triples of these equations, say between the first three and the last three, the cubic cones K_4 and K_1 with vertices A_4 and A_1 and the common generatrix A_1A_4 are obtained, along which they have the common tangent plane $b_2x_2 - b_3x_3$

$=0$. Hence they intersect in a residual septimic C_7, the locus of the point Q, Q'. This follows immediately by inspection of the equations

$$K_4 = b_1x_1(x_2^2 - x_3^2) + b_2x_2(x_3^2 - x_1^2) + b_3x_3(x_1^2 - x_2^2) = 0,$$

$$K_1 = b_2x_2(x_3^2 - x_4^2) + b_3x_3(x_4^2 - x_2^2) + b_4x_4(x_2^2 - x_3^2) = 0.$$

This can be verified by other methods of proof which for the sake of brevity shall be omitted.

(3) To prove that the genus of C_7 is three, project C_7 upon K_1 from a generic point P. The projection proper is a residual C_{14} of order $3\times7-7=14$. The cone K_4 cuts C_{14} in $3\times14-6=36$ points proper, because C_7 touches both K_1 and K_4 along A_1A_4 in three points which accounts for six (improper) points of intersection. The polar conic of P with respect to K_1 cuts C_7 outside of A_1 and A_4 in twelve points, so that altogether $36-12=24$ points of intersection are left which are projected into twelve double points of C_7', the projection of C_7 upon a generic plane. The genus p of C_7' and hence of C_7 is therefore $p=6\times5/2-12=3$.

Now every couple $P_i'P_i'$ on K or C_7 gives rise to a definite Δ_{16}-configuration. Hence we have our next theorem:

THEOREM 4. *On every cubic cone K of the cubic complex associated with the involutorial cubic transformation T_3 there exists an invariant septimic C_7 of genus three with ∞^1 Δ_{16}-configurations.*

It is interesting to note that the C_7 lies on two other cubic cones K_2 and K_3 with vertices at A_2 and A_3 by using the elimination process of λ, μ, 1 in the remaining possible ways, so that it may also be characterized by the property that *it lies on five cubic cones.*

(4) The investigation may be extended to any other S_r, $r>3$, but this would amount merely to a simple generalization of the preceding theory.

UNIVERSITY OF ILLINOIS

ON ADDITION CHAINS

ALFRED BRAUER

We consider a set $a_0=1<a_1<a_2<\cdots<a_r=n$ of integers such that every element a_ρ can be written as sum $a_\sigma+a_\tau$ of preceding elements of the set. Such sets of integers have been called "addition chains (Additionsketten) for n" by A. Scholz.† For example, for $n=666$,

$$1, 2, 4, 8, 16, 24, 40, 80, 160, 320, 640, 664, 666$$

forms an addition chain with $r=12$; the same holds for

$$1, 2, 3, 6, 9, 18, 27, 54, 81, 162, 324, 648, 666.$$

In any case, we must have $a_1=2$ and $a_2=3$ or 4.

By the length $l=l(n)$ of n, Scholz understands the smallest l for which there exists an addition chain $a_0, a_1, \cdots, a_l=n$.

The following question leads to addition chains: The least positive residue of $c^n \pmod m$ (c, m, n given integers) is to be formed using the smallest possible number of multiplications. Then $l(n)$ multiplications will always suffice.

A. Scholz published the following inequalities for $l(n)$ in the form of problems:

$$m+1 \leqq l(n) \leqq 2m \quad \text{for} \quad 2^m+1 \leqq n \leqq 2^{m+1}, \qquad m \geqq 1, \tag{1}$$

$$l(ab) \leqq l(a)+l(b). \tag{2}$$

In (1), we have $l(n)<2m$ whenever $m>2$; moreover,

$$l(2^{m+1}-1) \leqq m+l(m+1). \tag{3}$$

In connection with (3), Scholz surmises that (1) can be improved generally. He further raises the question of whether or not the inequality

$$1 \leqq \limsup_{n\to\infty} \frac{\log 2}{\log n} l(n) \leqq 2, \tag{4}$$

which easily follows from (1), can be improved.

It is easy to prove the formulas (1) and (2). I cannot decide whether (3) is always true. In the following, I will show that

$$l(2^{m+1}-1) \leqq m+l^*(m+1),$$

† Jahresbericht der deutschen Mathematiker-Vereinigung, class II, vol. 47 (1937), p. 41.

where $l^*(m+1)$ is the minimal length, not of all, but only of certain addition chains. Further, I will prove by elementary methods that for sufficiently large n

$$l(n) < \frac{\log n}{\log 2}\left\{1 + \frac{1}{\log\log n} + \frac{2\log 2}{(\log n)^{1-\log 2}}\right\}.$$

This is better than (1). It entails the following relation

$$\lim_{n\to\infty} \frac{\log 2}{\log n}\, l(n) = 1,$$

which, of course, is better than (4).

Let $a_0, a_1, \cdots, a_l = n$ be an addition chain for n, $2^m+1 \leqq n \leqq 2^{m+1}$. Then $a_\lambda \leqq 2a_{\lambda-1}$, $(\lambda = 1, 2, \cdots, l)$. Since $a_1 = 2$, we have $a_l \leqq 2^l$, $n \leqq 2^l$, $2^m+1 \leqq 2^l$, $m+1 \leqq l$. This proves the first half of (1).

To prove the second part of (1), $l(n) \leqq 2m$, suppose first that $2^m+1 \leqq n < 2^{m+1}$. We write n as a binary number

$$n = 2^{\nu_1} + 2^{\nu_2} + \cdots + 2^{\nu_k}, \quad \nu_1 < \nu_2 < \cdots < \nu_k.$$

We have here at most $m+1$ terms, $k \leqq m+1$, and $\nu_k = m$. We begin the addition chain with $a_0 = 1$, $a_1 = 2$, $a_2 = 4, \cdots, a_m = 2^m$, and take then

$$a_{m+1} = 2^m + 2^{\nu_1},\ a_{m+2} = 2^m + 2^{\nu_1} + 2^{\nu_2}, \cdots,$$
$$a_{m+k-1} = 2^m + 2^{\nu_1} + \cdots + 2^{\nu_{k-1}} = n.$$

This actually is an addition chain, and we see that $l(n) \leqq m+k-1 \leqq 2m$. The equality $l(n) = 2m$ is possible only if $k = m+1$,

$$n = 1 + 2 + 2^2 + \cdots + 2^m = 2^{m+1} - 1.$$

This case will be discussed in the last paragraph of this page.

For $n = 2^{m+1}$, we form the addition chain

$$1, 2, 2^2, \cdots, 2^{m+1}. \tag{5}$$

Here $l = m+1$, hence $l(2^{m+1}) = m+1 \leqq 2m$.

Let $1, a_1, a_2, \cdots, a_r = a$ be an addition chain for a with $r = l(a)$, and let $1, b_1, b_2, \cdots, b_s$ be one for b with $s = l(b)$. Then

$$1, a_1, \cdots, a_r, a_r b_1, a_r b_2, \cdots, a_r b_s$$

forms an addition chain for $a_r b_s = ab$, since $b_\rho = b_\sigma + b_\tau$ implies $a_r b_\rho = a_r b_\sigma + a_r b_\tau$. The number of terms after 1 in this chain is $r+s$; hence $l(ab) \leqq r+s = l(a)+l(b)$. This proves (2).

By a special addition chain for the number n we mean an addition chain for which, for all ρ, and for some σ

$$a_\rho = a_{\rho-1} + a_\sigma, \qquad 0 \leqq \sigma \leqq \rho - 1 \leqq l - 1,$$

holds. Let $l^*(n)$ be the minimal length of all special addition chains for n. Then $l(n) \leqq l^*(n)$. The chains used in the proof of the second part of (1) are special chains. Hence, it follows from this proof that $l^*(n) \leqq 2m$. The equality sign is here impossible except for $n = 2^{m+1} - 1$. In order to prove that $l(n) < 2m$ whenever $m > 2$, it suffices to show that

$$l(2^{m+1} - 1) \leqq l^*(2^{m+1} - 1) \leqq m + l^*(m + 1), \tag{6}$$

for 1, 2, 4, 5, 6, 7, $\cdots$, $m+1$ is a special chain of length $m-1$ for $m+1$, so $l^*(m+1) \leqq m-1$. Let

$$1 = a_0, a_1, \cdots, a_k = m + 1 \tag{7}$$

be a minimal special addition chain for $m+1$, $k = l^*(m+1)$. We form

$$2^{a_0} - 1 = 1, 2^{a_1} - 1 = 3, 2^{a_2} - 1, \cdots, 2^{a_k} - 1$$

and multiply $2^{a_\kappa} - 1$ successively $a_{\kappa+1} - a_\kappa$ times by 2, ($\kappa = 0, 1, 2, \cdots, k-1$). We then obtain

$$(a_1 - a_0) + (a_2 - a_1) + \cdots + (a_k - a_{k-1}) = m$$

further numbers. We thus obtain the integers

$$\begin{aligned} &1, 2, 2^{a_1} - 1, 2(2^{a_1} - 1), 2^2(2^{a_1} - 1), \cdots, 2^{a_2 - a_1}(2^{a_1} - 1), 2^{a_2} - 1, \\ &2(2^{a_2} - 1), \cdots, 2^{a_3 - a_2}(2^{a_2} - 1), \cdots, 2^{a_k - a_{k-1}}(2^{a_{k-1}} - 1), 2^{a_k} - 1. \end{aligned} \tag{8}$$

We state that these numbers form a special chain for $2^{a_k} - 1 = 2^{m+1} - 1$. This will be proved if we show that

$$2^{a_\kappa} - 1 - 2^{a_\kappa - a_{\kappa-1}}(2^{a_{\kappa-1}} - 1) = 2^{a_\kappa - a_{\kappa-1}} - 1$$

is an element of (8) for $\kappa = 2, 3, \cdots, k$. But this is true, since (7), as a special addition chain, contains $a_\kappa - a_{\kappa-1}$. The length of the chain (8) is $k + m = l^*(m+1) + m$, and this proves (6).

We show now that A. Scholz' conjecture, that (1) and (4) can be improved, is actually true. We prove the following theorem:

THEOREM. *If r is any positive and s any not negative integer, then*

$$l(n) \leqq (r + 1)s + 2^r - 2 \quad \textit{for} \quad 2^{rs} \leqq n < 2^{r(s+1)}. \tag{9}$$

PROOF. When $r = 1$, this follows from (1); we therefore take $r > 1$ and fixed. I state now that we can form an addition chain for n which contains at most $(r+1)s + 2^r - 2$ terms, and which begins with the terms $a_0 = 1, a_1 = 2, a_2 = 3, \cdots, a_{2^r-2} = 2^r - 1$. For $s = 0$ this is true because the integers 1, 2, $\cdots$, $2^r - 1$ form an addition chain for every

$n<2^r$ (the integers $n+1, n+2, \cdots, 2^r-1$ may be included in the chain). Assume now that the assertion is not true and take s to be the smallest value for which the statement does not hold for n with $2^{rs} \leqq n < 2^{r(s+1)}$. We divide n by 2^r:

$$n = a \cdot 2^r + b, \qquad 0 \leqq b < 2^r. \tag{10}$$

Then $2^{r(s-1)} \leqq a < 2^{rs}$, and since our statement is supposed to be true for $s-1$ instead of s, there exists an addition chain $a_0, a_1, a_2, \cdots, a_{a-1}, a_a = a$ for a which has at most $(r+1)(s-1)+2^r-2$ terms, and which starts with $1, 2, \cdots, 2^r-1$. Because of (10), this chain contains b for $b>0$. Then $a_0, a_1, \cdots, a_{a-1}, a, 2a, 2^2a, \cdots, 2^ra, 2^ra+b$ is an addition chain for n which contains the first 2^r-1 integers. The length equals at most

$$(r+1)(s-1) + 2^r - 2 + r + 1 = (r+1)s + 2^r - 2.$$

This gives the desired contradiction; therefore the statement holds for all values of s. The proof yields an easy method for constructing the addition chains.

From relation (9) it follows that $s \leqq (\log n)/r \cdot \log 2$; hence $l(n) \leqq (r+1)(\log n)/(r \log 2) + 2^r - 2$. If now $2^m \leqq n < 2^{m+1}$, this yields

$$l(n) \leqq \min_{r=1,2,\cdots,m} \left\{ \left(1 + \frac{1}{r}\right) \frac{\log n}{\log 2} + 2^r - 2 \right\}. \tag{11}$$

For instance, if we set $r = [\log \log n] + 1$ for $n \geqq 3$, then (11) gives

$$\begin{aligned} l(n) &< \left(1 + \frac{1}{\log \log n}\right) \frac{\log n}{\log 2} + 2^{\log \log n + 1} \\ &= \left(1 + \frac{1}{\log \log n}\right) \frac{\log n}{\log 2} + 2e^{\log \log n \cdot \log 2} \\ &= \frac{\log n}{\log 2} \left(1 + \frac{1}{\log \log n}\right) + 2(\log n)^{\log 2}, \end{aligned}$$

$$l(n) < \frac{\log n}{\log 2} \left\{ 1 + \frac{1}{\log \log n} + \frac{2 \log 2}{(\log n)^{1-\log 2}} \right\}. \tag{12}$$

This inequality can easily be improved since the expression between the braces in (11) takes on its minimum for $r^2 \cdot 2^r = (\log n)/(\log 2)^2$.

On the other hand, it follows from (1) that $l(n) \geqq m > (\log n)/(\log 2) - 1$. This, in connection with (12) yields $\lim_{n \to \infty} l(n)(\log 2)/(\log n) = 1$.

THE INSTITUTE FOR ADVANCED STUDY

A THEOREM ON MATRICES OVER A COMMUTATIVE RING

NEAL H. McCOY

1. **Introduction.** Let R be an arbitrary commutative ring with unit element 1, and $R[\lambda]$ the ring of polynomials in the indeterminate λ, with coefficients in R. If A is a matrix of order n, with elements in R, the set of all elements $g(\lambda)$ of $R[\lambda]$, such that $g(A)=0$, is an ideal which we shall call the *minimum ideal* of A. The element $f(\lambda)=|\lambda-A|$ of $R[\lambda]$ is the *characteristic function* of A, and the principal ideal $(f(\lambda))$ may be called the *characteristic ideal* of A.* In a recent note,† it was shown that the minimum ideal of a matrix can be characterized in a manner generalizing Frobenius' characterization of the minimum function of a matrix for the case in which the coefficient domain is a field.‡ It was also shown that, in $R[\lambda]$, the prime ideal divisors of the minimum ideal coincide with those of the characteristic ideal. If R is specialized to be an algebraically closed field, this result yields the familiar theorem to the effect that the distinct linear factors of the characteristic function of A coincide with the distinct linear factors of the minimum function of A. It is the primary purpose of the present note to generalize, in a similar way, the well known theorem of Frobenius concerning the characteristic roots of a polynomial in two or more commutative matrices—or, more precisely, an extension of this theorem which we shall now describe in some detail.

Let K denote an algebraically closed field, and let us say that the matrices A_i, $(i=1, 2, \cdots, m)$, with elements in K, have property P, if the characteristic roots of every scalar polynomial $f(A_1, A_2, \cdots, A_m)$, with coefficients in K, are all of the form $f(\lambda_1, \lambda_2, \cdots, \lambda_m)$ where λ_i is a characteristic root of A_i, $(i=1, 2, \cdots, m)$.

In a previous paper,§ the following statements were shown to be equivalent:

I. The matrices A_i, $(i=1, 2, \cdots, m)$, have property P.

* The terms *minimum ideal* and *characteristic ideal* are used merely to emphasize that they generalize the usual notions of minimum and characteristic functions, respectively.

† Neal H. McCoy, *Concerning matrices with elements in a commutative ring*, this Bulletin, vol. 45 (1939), pp. 280–284.

‡ For the classical theorems concerning the characteristic and minimum functions and related topics, see C. C. MacDuffee, *The Theory of Matrices*, chap. 2, or J. H. M. Wedderburn, *Lectures on Matrices*, chap. 2.

§ N. H. McCoy, *On the characteristic roots of matric polynomials*, this Bulletin, vol. 42 (1936), pp. 592–600. Hereafter this will be referred to as M.

II. All matrices $A_iA_j-A_jA_i$, $(i, j=1, 2, \cdots, m)$, are contained in the radical of the algebra of polynomials in the A's with coefficients in K.*

If the matrices A_i are commutative in pairs, clearly condition II is satisfied, and thus I is true. Hence Frobenius' theorem, which states that a set of commutative matrices always has property P, is a special case of the equivalence of I and II. Other interesting special cases of the general result stated above, or examples of matrices having property P, have been obtained by Bruton, Ingraham, Roth, and Williamson.†

If now the matrices A_i, $(i=1, 2, \cdots, m)$, have elements in an arbitrary commutative ring R with unit element, it is obvious that the above definition of property P no longer has any meaning. However, we shall give below a suitable definition of property P which is equivalent to the above, if R is specialized to be an algebraically closed field. The principal result of the present note is then a proof of the equivalence of I and II in this generalized sense.

2. **Preliminary remarks and notation.** We first recall a few properties of ideals which will be of importance in the sequel.‡ If $\mathfrak{a}$ is an ideal in the commutative ring R, the set of all elements of R, of which some finite power belongs to $\mathfrak{a}$, is an ideal called the *radical* of $\mathfrak{a}$. The radical of R is the radical of the null ideal, that is, the set of all nilpotent elements. Clearly $\mathfrak{a}$ and its radical have the same prime ideal divisors. A *minimal prime ideal divisor* of $\mathfrak{a}$ is one containing§ no other prime ideal divisor of $\mathfrak{a}$. *Each ideal* $\mathfrak{a}$ *has minimal prime ideal divisors, and the radical of* $\mathfrak{a}$ *is the intersection of all minimal prime ideal divisors of* $\mathfrak{a}$. Each prime ideal divisor of $\mathfrak{a}$ contains at least one minimal prime ideal divisor, so that, in fact, the radical of $\mathfrak{a}$ is the intersection of *all* prime ideal divisors of $\mathfrak{a}$.

Henceforth R will denote an arbitrary commutative ring with

* In other words, this may be described as follows, using the notation of the present paper. If S is the ring of polynomials in the A's over the ring R, and $\mathfrak{d}$ denotes the two-sided ideal in S generated by all elements $A_iA_j-A_jA_i$, then II states that all elements of $\mathfrak{d}$ are nilpotent.

† G. S. Bruton, *Certain aspects of the theory of equations for a pair of matrices*, this Bulletin, abstract 38-9-196; M. H. Ingraham, *A study of certain related pairs of square matrices*, ibid., abstract 38-9-197; W. E. Roth, *On the characteristic values of the matrix* $f(A, B)$, Transactions of this Society, vol. 39 (1936), pp. 234–243; J. Williamson, *The simultaneous reduction of two matrices to triangle form*, American Journal of Mathematics, vol. 57 (1935), pp. 281–293.

‡ For definitions and fundamental theorems, see W. Krull, *Idealtheorie*, particularly p. 9.

§ In the set-theoretic sense.

unit element, and A_i, $(i=1, 2, \cdots, m)$, fixed matrices of order n with elements in R. Let R_n denote the ring of all matrices of order n over R, and S the subring of R_n generated by the elements A_i, $(i=1, 2, \cdots, m)$, together with the unit element of R_n.* We shall denote by $\mathfrak{s}$ the two-sided ideal in S generated by the elements $A_{ij}=A_iA_j-A_jA_i$, $(i, j=1, 2, \cdots, m)$. Each element of $\mathfrak{s}$ is, therefore, expressible as a finite sum of terms of the form $gA_{ij}h$, where g and h are elements of S.

If y is any element of S, let $\bar{y}$ be the corresponding element of the ring $S/\mathfrak{s}$ under the homomorphic correspondence $S \rightarrow S/\mathfrak{s}$. We now introduce the ring $R'=R[x_1, x_2, \cdots, x_m]$ of polynomials in the commutative indeterminates $x_1, x_2, \cdots, x_m$ over R. To each polynomial $g(A)=g(A_1, A_2, \cdots, A_m)$ in the A's we may, therefore, make correspond the element $g(x)=g(x_1, x_2, \cdots, x_m)$ of R', obtained by formally replacing A_i by x_i, $(i=1, 2, \cdots, m)$.† For example, if $g(A)=A_1A_2A_1$, then $g(x)=x_1^2x_2$. Since $A_iA_j \equiv A_jA_i$ $(\mathfrak{s})$, it follows that multiplication is commutative in $S/\mathfrak{s}$. Thus the correspondence

$$f(x) = \sum a_{i_1 i_2 \cdots i_m} x_1^{i_1} x_2^{i_2} \cdots x_m^{i_m} \rightarrow \sum \bar{a}_{i_1 i_2 \cdots i_m} \overline{A}_1^{i_1} \overline{A}_2^{i_2} \cdots \overline{A}_m^{i_m},$$

which we may abbreviate in the form $f(x) \rightarrow \bar{f}(\overline{A})$, is a homomorphism between R' and $S/\mathfrak{s}$, and thus

$$S/\mathfrak{s} \cong R'/\mathfrak{m},$$

where $\mathfrak{m}$ is the ideal in R' of all elements $f(x)$ such that $\bar{f}(\overline{A})=0$. Clearly, $\mathfrak{m}$ contains the minimum ideal of each of the individual matrices A_i, $(i=1, 2, \cdots, m)$.

Let $\mathfrak{p}_\alpha$‡ denote an arbitrary minimal prime ideal divisor of $\mathfrak{m}$, and $\mathfrak{r}$ the radical of $\mathfrak{m}$, so that $\mathfrak{r}$ is the intersection of all $\mathfrak{p}_\alpha$. With each $\mathfrak{p}_\alpha$ we may associate, by means of a given polynomial $f(A)$, a prime ideal $\mathfrak{p}_\alpha'$ in $R[\lambda]$, whose elements are the polynomials $t(\lambda)$, such that $t[f(x)] \equiv 0$ $(\mathfrak{p}_\alpha)$. For convenience, we may indicate the intersection of all $\mathfrak{p}_\alpha'$ by $\mathfrak{k}$. We remark that $h[f(x)] \equiv 0$ $(\mathfrak{r})$, if and only if $h(\lambda) \equiv 0$ $(\mathfrak{k})$.

* Elements of S are, therefore, expressible as polynomials in the A_i, that is, as finite sums of terms of the form $aA_{i_1}A_{i_2} \cdots A_{i_k}$, where a is in R and each A_{i_j} is some one of the matrices $A_1, A_2, \cdots, A_m$. It may happen that no A_i appears in a term, in which case the term will be simply of the form a, as we shall not distinguish between the unit element of R and the unit element of R_n.

† This use of the symbols "A" and "x" will cause no confusion, as they do not appear without subscripts in any other connection. However, λ will always denote a single indeterminate.

‡ The use of a subscript is not meant to imply that the number of minimal prime ideal divisors is finite or even enumerable. The range of α may be an arbitrary set.

Let $\mathfrak{n}$ be the minimum ideal of the matrix $f(A)$, that is, the set of all polynomials $g(\lambda)$ such that $g[f(A)]=0$. We now make the following definition:

DEFINITION. *The matrices A_i, $(i=1, 2, \cdots, m)$, with elements in R are said to have property P, if for every polynomial $f(A)$, the radical of $\mathfrak{n}$ is $\mathfrak{k}$.*

Before proceeding, we pause to point out briefly the meaning of property P, if R is specialized to be an algebraically closed field. Let $f(A)$ be a given polynomial in the A_i. Since $\mathfrak{m}$ contains the characteristic function $g_j(x_j)$ of A_j, $(j=1, 2, \cdots, m)$, it follows* that each prime ideal divisor of $\mathfrak{m}$ is necessarily of the form

$$\mathfrak{p}_\alpha = (x_1 - \lambda_1^{(\alpha)}, x_2 - \lambda_2^{(\alpha)}, \cdots, x_m - \lambda_m^{(\alpha)}),$$

where $\lambda_j^{(\alpha)}$ is a characteristic root of A_j, $(j=1, 2, \cdots, m)$, and further that each $\mathfrak{p}_\alpha$ is minimal. By a Taylor's series expansion we see at once that

$$\mathfrak{p}_\alpha' = (\lambda - f(\lambda_1^{(\alpha)}, \lambda_2^{(\alpha)}, \cdots, \lambda_m^{(\alpha)}),$$

and that the prime ideal divisors of $\mathfrak{k}$ are precisely these $\mathfrak{p}_\alpha'$. Now the prime ideal divisors of $\mathfrak{n}$ are of the form $\lambda - a_j$, where a_j varies over the distinct characteristic roots of $f(A)$. Thus, if the given matrices have property P according to the definition above, the radical of $\mathfrak{n}$ is $\mathfrak{k}$ and, therefore, the prime ideal divisors of $\mathfrak{n}$ coincide with the prime ideal divisors of $\mathfrak{k}$. This means that the characteristic roots of $f(A_1, A_2, \cdots, A_m)$ are all of the form $f(\lambda_1^{(\alpha)}, \lambda_2^{(\alpha)}, \cdots, \lambda_m^{(\alpha)})$, and thus that the matrices have property P as defined in the introduction. Conversely, it is not difficult to show,† although we shall omit the proof, that if the matrices have property P as defined in the introduction, they also have property P as defined here. Thus, if R is an algebraically closed field, the two definitions are equivalent.

3. **The main theorem.** We now prove the following theorem which is the principal result of this note:

THEOREM. *Let A_i, $(i=1, 2, \cdots, m)$, be matrices of order n with elements in an arbitrary commutative ring R with unit element, and denote by S the ring of polynomials in the A_i with coefficients in R. If $\mathfrak{s}$ is the two-sided ideal in S generated by the matrices $A_iA_j - A_jA_i$,*

* For theorems on polynomial ideals see van der Waerden, *Moderne Algebra*, vol. 2.

† Cf. M, pp. 598–599.

$(i, j = 1, 2, \cdots, m)$, then a necessary and sufficient condition that the given matrices have property P is that all elements of $\mathfrak{s}$ be nilpotent.

First, we assume that all elements of $\mathfrak{s}$ are nilpotent. If $f(A)$ is an arbitrary polynomial in the A_i, we shall show that $\mathfrak{k}$ is the radical of $\mathfrak{n}$. Let $g(\lambda)$ be any element of $\mathfrak{k}$. Then it follows that $g[f(x)] \equiv 0$ $(\mathfrak{r})$, where $\mathfrak{r}$ is the radical of $\mathfrak{m}$. Thus, for some positive integer k, $\{g[f(x)]\}^k \equiv 0$ $(\mathfrak{m})$. This means, however, that $\{g[f(A)]\}^k \equiv 0$ $(\mathfrak{s})$, and, since all elements of $\mathfrak{s}$ are nilpotent, there exists a positive integer l such that $\{g[f(A)]\}^{kl} = 0$. This implies that $[g(\lambda)]^{kl} \equiv 0$ $(\mathfrak{n})$, that is, that $g(\lambda)$ is in the radical of $\mathfrak{n}$. Thus $\mathfrak{k}$ is contained in the radical of $\mathfrak{n}$.

Now let $h(\lambda)$ be an arbitrary element of the radical of $\mathfrak{n}$, that is, $[h(\lambda)]^\alpha \equiv 0$ $(\mathfrak{n})$, for some positive integer α. Then $\{h[f(A)]\}^\alpha = 0$, from which it follows that $\{h[\bar{f}(\bar{A})]\}^\alpha = 0$, and thus that

$$\{h[f(x)]\}^\alpha \equiv 0 \ (\mathfrak{m}).$$

This means that $h[f(x)] \equiv 0$ $(\mathfrak{r})$, and this, in turn, implies that $h(\lambda) \equiv 0$ $(\mathfrak{k})$. We have therefore shown that if all elements of $\mathfrak{s}$ are nilpotent, the radical of $\mathfrak{n}$ is $\mathfrak{k}$.

Conversely, let us now assume that for every polynomial $f(A)$, the radical of $\mathfrak{n}$ is $\mathfrak{k}$. Select any element of $\mathfrak{s}$ and write it, in any way, in the form of a sum of terms

$$F(A)(A_iA_j - A_jA_i)G(A),$$

and denote by $f(A)$ the polynomial thus obtained. Then, clearly, $f(x) = 0$, and thus $f(x) \equiv 0$ $(\mathfrak{r})$. Hence $\lambda \equiv 0$ $(\mathfrak{k})$ and therefore, by hypothesis, there exists a positive integer β such that $\lambda^\beta \equiv 0$ $(\mathfrak{n})$. It follows that $[f(A)]^\beta = 0$, and thus that $f(A)$ is nilpotent. The theorem is therefore established.

SMITH COLLEGE

AN APPLICATION OF E. H. MOORE'S DETERMINANT OF A HERMITIAN MATRIX*

N. JACOBSON

E. H. Moore defined a determinant for any hermitian matrix with elements in a "number system of type B."† In more descriptive language, a system Φ of this type may be characterized as a quasi-field of characteristic not equal to 2 in which there is defined an involutorial anti-automorphism or *involution* $a \to \bar{a}$:

$$\overline{a+b} = \bar{a} + \bar{b}, \qquad \overline{ab} = \bar{b}\bar{a}, \qquad \bar{\bar{a}} = a,$$

such that the symmetric elements ($\bar{a} = a$) are contained in the center. It follows readily that Φ is either commutative with $\bar{a} \equiv a$, a quadratic field over the field of symmetric elements, or a generalized quaternion algebra over this field. An examination of Moore's theory of determinants shows that it is entirely integral, and hence it is valid if Φ is any ring with an identity in which there is a unique element $1/2$ such that $2(1/2) \equiv 1/2+1/2 = 1$ and which has an involution $a \to \bar{a}$ whose symmetric elements are in the center Γ of Φ.

The uniqueness of $1/2$ implies its symmetry. If $2a \equiv a+a = 0$, then $0 = (a+a)/2 = a/2+a/2 = (1/2+1/2)a = a$. Let Σ and P respectively denote the sets of symmetric and of skew elements ($\bar{a} = -a$) of Φ. Then Σ and P are subgroups under the operation $+$. If $b \in \Sigma \cap \mathrm{P}$, $b = -b$, $2b = 0$, and hence $b = 0$. For any a we have

$$a = \tfrac{1}{2}(a + \bar{a}) + \tfrac{1}{2}(a - \bar{a}) = Sa + Va,$$

where $Sa \in \Sigma$, $Va \in \mathrm{P}$. Thus the additive group of Φ is a direct sum of Σ and P. We call Sa and Va respectively the scalar and the vector parts of a. Now Σ is a subring of Γ, and P is closed under multiplication by elements in Σ and under commutation $[u, v] = uv - vu$. Hence, for any two elements a, b, $[a, b] = [Va, Vb] \in \mathrm{P}$. Thus $S[a, b] = 0$ and since, in general, $S(a+b) = Sa+Sb$, $Sab = Sba$. Moreover, $a\bar{a}$ and $\bar{a}a$ are symmetric, $a\bar{a} - \bar{a}a$ skew. Hence $a\bar{a} = \bar{a}a$. As usual we call this element, the norm of a, Na and note that $Nab = (Na)(Nb)$. Any element a satisfies a quadratic equation with coefficients in Σ, namely,

$$x^2 - (2Sa)x + Na = 0.$$

* Presented to the Society, February 25, 1939.

† See Moore and Barnard, *General Analysis* I, American Philosophical Society Publication, chap. 2. We refer to this volume as M-B.

Thus all the properties noted in M-B, page 103, which do not involve inverses hold in the present case.

We consider the ring Φ_n of n-rowed square matrices with elements in Φ. The correspondence $A=(a_{ij})\to\overline{A}'=(b_{ij})$, where $b_{ij}=\bar{a}_{ji}$ is an involution in Φ_n. As is customary, we call the symmetric elements $(\overline{A}'=A)$ under this involution hermitian and we denote their totality as H. If A, B are in H, X arbitrary, then $A\pm B$, $AB+BA$ and $\overline{X}'AX$ are in H.

Suppose $A=(a_{ij})$ is hermitian. If σ is a set of distinct indices in the range $1,\cdots,n$ and f is in σ, we define

$$s_\sigma=\sum(-1)^r a_{fh_1}a_{h_1h_2}\cdots a_{h_{r-1}h_r}a_{h_rf},$$

summed on all permutations of the indices h_i of σ which are not equal to f. It can be shown that s_σ is in Σ and does not depend on the particular index f selected in σ.* Moore's determinant is defined as

$$\det_M A=\sum s_{\sigma_1}s_{\sigma_2}\cdots s_{\sigma_k},$$

summed on all partitions of $1,\cdots,n$ into disjoint sets σ_i. The proofs in M-B, pages 110–123, are valid for the present case. We require in particular the existence of the adjoint, that is, a certain hermitian matrix $\mathrm{adj}_M A$ whose elements are polynomials of $(n-1)$st degree in those of A, such that

$$A(\mathrm{adj}_M A)=(\mathrm{adj}_M A)A=(\det_M A)1,$$

where 1 is the identity matrix. It follows from this as usual that A satisfies its characteristic equation $\det_M(x1-A)=0$ of nth degree in x with coefficients in Σ that are polynomials in the elements of the matrix A.†

We now apply this result to ordinary matrix theory. Let Σ be any field of characteristic not equal to 2, Σ_{2n} the $2n$-rowed matrix ring over Σ and Q the matrix

$$Q=\begin{pmatrix} q & & & \\ & q & & \\ & & \ddots & \\ & & & q \end{pmatrix},\qquad q=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The correspondence $A\to Q^{-1}A'Q$ is an involution in Σ_{2n} and, if

* M-B, p. 114.

† See, for example, Wedderburn, *Lectures on Matrices*, p. 23, or Albert, *Modern Higher Algebra*, p. 78. Compare also M-B, p. 128.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

where $a_{ij} \in \Sigma_2$,

$$Q^{-1}A'Q = \begin{pmatrix} \bar{a}_{11} & \bar{a}_{21} & \cdots & \bar{a}_{n1} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{nn} \end{pmatrix},$$

where, in general,

$$\bar{a} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} \quad \text{if} \quad a = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

The correspondence $a \to \bar{a}$ is an involution in $\Phi = \Sigma_2$ satisfying the above conditions, and $A \to Q^{-1}A'Q$ is equivalent to $A = (a_{ij}) \to \overline{A}'$. Hence, if A is symmetric in the sense that $Q^{-1}A'Q = A$ ($\overline{A}' = A$), the above result shows that A satisfies an equation of the nth degree in x with coefficients which are polynomials in the elements $\alpha, \beta, \cdots$.

Now suppose R is any nonsingular matrix such that $R' = -R$. If B satisfies $R^{-1}B'R = B$ and V is any matrix,

$$(V'RV)^{-1}(V^{-1}BV)'(V'RV) = V^{-1}BV.$$

We may choose V so that $V'RV = Q$ as above.* Thus $A = V^{-1}BV$ is symmetric with respect to the involution $A \to \overline{A}'$. Hence B satisfies an equation $\phi(x; \alpha, \beta, \cdots)$ of the nth degree with coefficients polynomials in the elements of B. Let $f(x; \alpha, \beta, \cdots)$ be the ordinary characteristic polynomial and $\psi(x; \alpha, \beta, \cdots)$ the minimum polynomial obtained when we regard the elements $\alpha, \beta, \cdots$ in the general matrix satisfying our condition as indeterminates.† As is well known, $f(x)$ and $\phi(x)$ are divisible by $\psi(x)$ and $f(x)$ and $\psi(x)$ have the same irreducible factors.

We now choose W so that

$$W'RW = S = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix},$$

where 1_n is the identity in Σ_n. Then $C = W^{-1}BW$ satisfies $S^{-1}C'S = C$. If

* Wedderburn, p. 96 or Albert, p. 108, loc. cit.

† The conditions on $A = V^{-1}BV$ are $a_{ii} = \alpha_{ii}1_2$ and $a_{ij} = \bar{a}_{ij}$, $i \neq j$. We regard α_{ii} and the elements of a_{ij}, $i < j$ as indeterminates. If we replace Σ by the polynomial ring obtained by adjoining these indeterminates to Σ, we obtain as before $\phi(B) = 0$.

$$C = \begin{pmatrix} L & M \\ N & P \end{pmatrix},$$

L, M, and so on, being in Σ_n, it follows that $M' = -M$, $N' = -N$, $P = L'$ and conversely any matrix of this form satisfies the condition. If we specialize the indeterminates so that $M = N = 0$, we obtain a matrix with irreducible minimum polynomial of degree n. It follows that $\psi(x)$ for the general matrix has degree not less than n and hence $\psi(x) = \phi(x)$ and $f(x) = [\phi(x)]^2$.

THEOREM. *If B is a matrix of $2n$ rows and columns with elements in a field of characteristic not equal to 2 such that $R^{-1}B'R = B$, where R is any non-singular skew symmetric matrix, then the characteristic polynomial $f(x)$ of B has the form $[\phi(x)]^2$, where the coefficients of $\phi(x)$ are polynomials in the elements of B and $\phi(B) = 0$. If the elements of the general matrix B are regarded as indeterminates, then $\phi(x)$ is irreducible.*

UNIVERSITY OF NORTH CAROLINA

THE EULER-MACLAURIN SUMMATION FORMULA*

TOMLINSON FORT

In this paper an extension of the classical Euler-Maclaurin summation formula† is made to multiple sums. Bernoulli polynomials and numbers of higher order as defined by Nörlund‡ enter into the formula and Bernoulli numbers of negative order enter into the proof. Nörlund obtains§ a formula for $\phi(x+\omega)$ in terms of Bernoulli numbers of higher order, and this is called by him an extension of the Euler-Maclaurin formula. His formula permits the ready building up of a simple sum. This is not true, however, of a multiple sum. Steffensen‖ calls attention to the fact that a multiple sum can be reduced by summation by parts to a simple sum and the Euler-Maclaurin formula for the simple sum used. However, the function to be summed is changed by his suggested transformation and he develops no general formula, nor does he suggest the use of Bernoulli numbers of higher order.

The formula developed in the present paper is equally as easy of

* Presented to the Society, October 29, 1938.

† See, for example, D. Seliwanoff, *Lehrbuch der Differenzenrechnung*, p. 48, where it is called the Euler sum formula. It is sometimes also called the Maclaurin sum formula. See W. B. Ford, *Studies on Divergent Series and Summability*.

‡ N. E. Nörlund, *Differenzenrechnung*, p. 129.

§ Loc. cit., p. 160.

‖ J. F. Steffensen, *Interpolation*, p. 136.

application as the classical formula. It is applied to one example. Throughout the paper use is freely made of the properties which Nörlund has developed of $B_n^{(l)}(x)$ and $B_n^{(l)}$, the Bernoulli polynomials and numbers of order l and degree n.

1. **The formula.** Let

$$
\begin{aligned}
f(x) &= P_0(n+k)(n+k-1)\cdots nx^{n-1} \\
&\quad + P_1(n+k-1)\cdots(n-1)x^{n-2}+\cdots \\
&\quad + P_{n-1}(k+1)\cdots 2\cdot 1,
\end{aligned} \tag{1}
$$

where the P's are constants, and let ω be a difference interval. Then*

$$
\begin{aligned}
\Phi_1(f(x)) &\equiv \omega \sum_{x_1=0}^{x-\omega} f(x_1) = P_0(n+k)(n+k-1)\cdots(n+1)[B_n^{(1)}(x) \\
&\quad - B_n^{(1)}] + \cdots + P_{n-1}(k+1)\cdots 2[B_1^{(1)}(x) - B_1^{(1)}], \\
\Phi_2(f(x)) &\equiv \omega^2 \sum_{x_1=0}^{x-\omega}\sum_{x_2=0}^{x_1-\omega} f(x_2) \\
&= P_0(n+k)\cdots(n+2)[B_{n+1}^{(2)}(x) - B_{n+1}^{(2)}] + \cdots \\
&\quad + P_{n-1}(k+1)\cdots 3[B_2^{(2)}(x) - B_2^{(2)}] - [P_0(n+k) \\
&\quad \cdots (n+1)B_n^{(1)} + \cdots + P_{n-1}(k+1)\cdots 2B_1^{(1)}]x, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \\
\Phi_{k+1}(f(x)) &\equiv \omega^{k+1} \sum_{x_1=0}^{x-\omega}\cdots\sum_{x_{k+1}=0}^{x_k-\omega} f(x_{k+1}) = P_0[B_{n+k}^{(k+1)}(x) - B_{n+k}^{(k+1)}] + \cdots \\
&\quad + P_{n-1}[B_{k+1}^{(k+1)}(x) - B_{k+1}^{(k+1)}] \\
&\quad - [P_0(n+k)B_{n+k-1}^{(k)} + \cdots + P_{n-1}(k+1)B_k^{(k)}]x \\
&\quad - [P_0(n+k)(n+k-1)B_{n+k-2}^{(k-1)} + \cdots \\
&\quad + P_{n-1}(k+1)kB_{k-1}^{(k-1)}]\frac{x(x-\omega)}{2!} - \cdots \\
&\quad - [P_0(n+k)\cdots(n+1)B_n^{(1)} + \cdots \\
&\quad + P_{n-1}(k+1)\cdots 2B_1^{(1)}]\cdot\frac{x(x-\omega)\cdots(x-(k-2)\omega)}{(k-1)!}.
\end{aligned} \tag{2}
$$

* Notation throughout the paper is $\sum_{x=a}^{x-\omega} f(x) = f(a)+f(a+\omega)+\cdots+f(x-\omega)$. We also use the convention $\sum_{x_1=a}^{x-\omega}\sum_{x_2=a}^{x_1-\omega} f(x_2) = 0$ if $x-\omega \leqq a$. It is well to remark that in case $f(x) \equiv f(x_1)$ is a polynomial and $\sum_{x_1=0}^{x-\omega} f_{n-1}(x_1) = f_n(x)$, then $f_n(0) = f_n(\omega) = \cdots = f_n((n-1)\omega) = 0$. Consequently, $\sum_{x_1=0}^{x-\omega}\cdots\sum_{x_n=0}^{x_{n-1}-\omega} f(x_n)$ can be evaluated by successive summations without calling into play any convention.

Rearrange the terms in (2) and we have

$$\begin{aligned}\Phi_{k+1}(f(x)) = {} & B_0^{(k+1)}[P_0x^{n+k} + P_1x^{n+k-1} + \cdots + P_{n-1}x^{k+1}] \\ & + B_1^{(k+1)}[P_0(n+k)x^{n+k-1} + \cdots + P_{n-1}(k+1)x^k] \\ & + B_2^{(k+1)}\left[P_0\frac{(n+k)(n+k-1)}{2!}x^{n+k-2} + \cdots\right. \\ & \left. + P_{n-1}\frac{(k+1)k}{2!}x^{k-1}\right] + \cdots + B_{n+k}^{(k+1)}[P_0] \\ & - [P_{n-1}B_{k+1}^{(k+1)} + \cdots + P_0B_{n+k}^{(k+1)}] \\ & - [P_{n-1}(k+1)B_k^{(k)} + \cdots + P_0(n+k)B_{n+k-1}^{(k)}]x - \cdots \\ & - [P_{n-1}(k+1)\cdots 2B_1^{(1)} + \cdots \\ & + P_0(n+k)\cdots(n+1)B_n^{(1)}] \\ & \cdot\frac{x(x-\omega)\cdots(x-(k-1)\omega)}{k!}.\end{aligned}$$

From this, by inspection,

$$\begin{aligned}\Phi_{k+1}(f(x)) = {} & B_0^{(k+1)}\int_0^x\int_0^{x_1}\cdots\int_0^{x_k} f(x_{k+1})dx_1dx_2\cdots dx_{k+1} \\ & + B_1^{(k+1)}\int_0^x\cdots\int_0^{x_{k-1}} f(x_k)dx_1\cdots dx_k \\ & + \frac{1}{2!}B_2^{(k+1)}\int_0^x\cdots\int_0^{x_{k-2}} f(x_{k-1})dx_1\cdots dx_{k-1} + \cdots \\ & + \frac{1}{k!}B_k^{(k+1)}\int_0^x f(x_1)dx_1 + \frac{1}{(k+1)!}B_{k+1}^{(k+1)}f(x) + \cdots \\ (3)\qquad & + \frac{1}{(n+k)!}B_{n+k}^{(k+1)}f^{(n-1)}(x) \\ & - \left[B_{k+1}^{(k+1)}\frac{f(0)}{(k+1)!} + \cdots + B_{n+k}^{(k+1)}\frac{f^{(n-1)}(0)}{(n+k)!}\right] \\ & - \left[B_k^{(k)}\frac{f(0)}{k!} + \cdots + B_{n+k-1}^{(k)}\frac{f^{(n-1)}(0)}{(n+k-1)!}\right]x - \cdots \\ & - \left[B_1^{(1)}\frac{f(0)}{1!} + \cdots + B_n^{(1)}\frac{f^{(n-1)}(0)}{n!}\right] \\ & \cdot\frac{x(x-\omega)\cdots(x-(k-1)\omega)}{k!}.\end{aligned}$$

This is the desired generalization of the Euler-Maclaurin formula. As developed for (1), it is a polynomial identity.

We now assume $f(x)$ not a polynomial and proceed to an examination of the remainder when the right-hand member of (3) replaces $\Phi_{k+1}(f(x))$. We assume x, ω and $f(x)$ real and $x \geqq (k+1)\omega$, also that all derivatives, which enter, exist. Denote the right-hand member of (3) by Ψ_{k+1}. Let $R_n(x) = \Phi_{k+1} - \Psi_{k+1}$. We now use a generalization* of Taylor's formula and obtain

$$
\begin{aligned}
-\Delta^{k+1}R_n(x) = -\omega^{k+1}f(x) + B_0^{k+1}\Big[& f(x) + \frac{\Delta^{k+1}0^{k+2}}{(k+2)!}f'(x) \\
& + \frac{\Delta^{k+1}0^{k+3}}{(k+3)!}f''(x) + \cdots \\
& + \frac{\Delta^{k+1}0^{n+k}}{(n+k)!}f^{(n-1)}(x) + R_0^{(k+1)}\Big] \\
& + B_1^{(k+1)}\Big[f'(x) + \frac{\Delta^{k+1}0^{k+2}}{(k+2)!}f''(x) + \cdots \\
& + (\ )f^{(n-1)}(x) + R_1^{(k+1)}\Big] \\
& + \frac{1}{2!}B_2^{(k+1)}\Big[f''(x) + \cdots + (\ )f^{(n-1)}(x) + R_2^{(k+1)}\Big] \\
& + \cdots + \frac{1}{(n+k)!}B_{n+k}^{(k+1)}[R_{n+k}^{(k+1)}].
\end{aligned} \tag{4}
$$

Here†

$$
\begin{aligned}
R_j^{(k+1)} = & \frac{1}{(n+k-j)!}\sum_{i=0}^{k}(-1)^i \\
& \cdot \int_0^{(k+1-i)\omega} C_{k+1,k+1-i}\, t^{n+k-j} f^{(n)}(x + (k+1-i)\omega - t)\,dt
\end{aligned} \tag{5}
$$

We note that $B_0^{(i)} = 1$, $i = 1, 2, \cdots$. We also now note the follow-

* George Boole, *Calculus of Finite Differences*, p. 24. In the present paper in calculating $\Delta^n 0^n$ the difference interval is ω.

† This is readily obtained if we write

$$\Delta^{k+1}Q(x) = \sum_{i=0}^{k}(-1)^i C_{k+1,k+1-i}[Q(x + (k+1-i)\omega) - Q(x)]$$

and use the integral form for the remainder when $Q(x+(k+1-i)\omega) - Q(x)$ is developed by Taylor's formula. Here, as throughout the paper, $C_{n,m}$ denotes a binomial coefficient.

ing relation between the numbers* $\Delta^n 0^n$ and the Bernoulli numbers of negative order as defined by Nörlund with all difference intervals equal:

$$\omega^n B_\nu^{(-n)} = \frac{\nu!}{(\nu+n)!}\Delta^n 0^{\nu+n}.$$

Substituting in (4) and rearranging terms we have:

$$-\Delta^{k+1}R_n(x) = \omega^{k+1}\sum_{j=1}^{n}\frac{f^j(x)}{j!}\sum_{i=0}^{j}C_{j,i}B_i^{(k+1)}B_{j-i}^{-(k+1)} + \sum_{j=0}^{n+k}\frac{1}{j!}B_j^{(k+1)}R_j^{(k+1)}.$$

But the first sum in the right-hand member is zero.† Hence we obtain

$$\Delta^{k+1}R_n(x) = -\sum_{j=0}^{n+k}\frac{1}{j!}B_j^{(k+1)}R_j^{(k+1)}.$$

Substituting for $R_j^{(k+1)}$ its value from (5), we have

$$\Delta^{k+1}R_n(x) = \frac{-1}{(n+k)!}\sum_{i=0}^{k}\int_0^{(k+1-i)\omega}(-1)^i C_{k+1,k+1-i}B_{n+k}^{(k+1)}(t) \cdot f^{(n)}(x+(k+1-i)\omega-t)dt. \tag{6}$$

Moreover,

$$R_n(0) = \Delta R_n(0) = \cdots = \Delta^k R_n(0) = 0.$$

To show this, refer to the definition of R_n as $\Phi_{k+1}-\Psi_{k+1}$ and use (3). It is immediate that $R_n(0)=0$. To show that $\Delta R_n(0), \cdots, \Delta^k R_n(0)$ are also zero, let $f(x)$ be a polynomial of arbitrary degree. Develop $\Delta^i R_n(x)$, $i=1, \cdots, k$, by the extended Taylor's series as given by Boole.‡ All series are finite. Let $x=0$. We have a polynomial identity in $f(0), f'(0), \cdots$. Since $f(0), f'(0), \cdots$ are arbitrary, their coefficients must be zero. These are expressions in the Bernoulli numbers that are independent of the function considered. Hence,

$$R_n(x) = \sum_{x_1=0}^{x-\omega}\sum_{x_2=0}^{x_1-\omega}\cdots\sum_{x_{k+1}=0}^{x_k-\omega}\Delta^{k+1}R_n(x_{k+1}), \tag{7}$$

* Boole, loc. cit., p. 28.
† Nörlund, loc. cit., p. 141.
‡ Loc. cit., p. 24

where $\Delta^{k+1}R_n(x)$ is given by (6). This form for $R_n(x)$ reduces to the classical Jacobi form* in case $k=0$.

The transformation of the interval $(0, x)$ into (a, y) in (3) and (7) by means of the transformation $y=a+x$ is immediate.

2. **An example.** Consider $\Phi_{k+1}(1/(x+1)^\alpha)$, with $\omega=1$, $\alpha>0$. Then

$$R_n(x) = -\frac{1}{(n+k)!}\sum_{i=0}^{k}\int_0^{k+1-i}\sum_{x_1=0}^{x-1}\cdots\sum_{x_{k+1}=0}^{x_k-1}L(x, t, i, n)dt, \tag{8}$$

where

$$L \equiv (-1)^{i+n}C_{k+1,k+1-i}B_{n+k}^{(k+1)}(t)\alpha(\alpha+1) \cdots (\alpha+n-1)(x+k+2-i-t)^{-(\alpha+n)}. \tag{9}$$

Adopting the notation $x^{(l)}=x(x-1)\cdots(x-l+1)$, it is readily shown by mathematical induction on k that

$$\sum_{x_1=0}^{x-1}\cdots\sum_{x_{k+1}=0}^{x_k-1}L = A_0x^{(k)} + A_1x^{(k-1)} + \cdots + A_k + M, \tag{10}$$

where the A's are independent of x and

$$M = (-1)^{k+1}\sum_{x_1=x}^{\infty}\cdots\sum_{x_{k+1}=x_k}^{\infty}L.$$

From this

$$|M| \leqq \int_{x-1}^{\infty}\cdots\int_{x_k-1}^{\infty}\frac{L'}{(x_{k+1}+1)^{\alpha+n}}dx_1\cdots dx_{k+1} \leqq \frac{L''}{(x+1)^{\alpha+n-k-1}},$$

where L' and L'' are certain positive constants. Here, in order to assure the convergence of all series under discussion, it is necessary to assume $n>k+1-\alpha$. We next integrate (10) as indicated in (8), and let

$$M' = \int_0^{k+1-i}Mdt.$$

Then $|M'| \leqq (k+1-i)L''/(x+1)^{\alpha+n-k-1}$. Let

$$M'' = \frac{-1}{(n+k)!}\sum_{i=0}^{k}M'.$$

Then, $|M''| \leqq L'''/(x+1)^{\alpha+n-k-1}$, where L''' is a constant. Let

* Seliwanoff, loc. cit., p. 50, formula (4).

$$D_j = \frac{-1}{(n+k)!} \sum_{i=0}^{k} \int_0^{k+1-i} A_j dt,$$

$$C_j = D_j - \frac{1}{(k-j)!}\left[B_{j+1}^{(j+1)} \frac{f(0)}{(j+1)!} + \cdots + B_{n+j}^{(j+1)} \frac{f^{(n-1)}(0)}{(n+j)!}\right].$$

Then

$$\begin{aligned}\Phi_{k+1}\left(\frac{1}{(x+1)^\alpha}\right) &= C_0 x^{(k)} + C_1 x^{(k-1)} + \cdots + C_k \\ &\quad + B_0^{(k+1)} \int_0^x \cdots \int_0^{x_k} \frac{1}{(x_{k+1}+1)^\alpha} dx_1 \cdots dx_{k+1} \\ &\quad + \cdots + \frac{1}{k!} B_k^{(k+1)} \int_0^x \frac{1}{(x_1+1)^\alpha} dx_1 \\ &\quad + \frac{1}{(k+1)!} B_{k+1}^{(k+1)} \frac{1}{(x+1)^\alpha} + \cdots \\ &\quad + \frac{1}{(n+k)!} B_{n+k}^{(k+1)} \frac{\alpha(\alpha+1)\cdots(\alpha+n-2)}{(x+1)^{\alpha+n-1}} + M''.\end{aligned} \tag{11}$$

We readily show that the C's are independent of n. Equate right-hand members of (11) for different values of n. Cancel integral terms. Divide through by x^k and allow x to become infinite. Cancel $C_0 x^k$ from both members and repeat. The C's can be calculated for particular values of k and α. In some instances the calculation may be precise, in others necessarily approximate. For example, if $\alpha=1$ and $k=0$, C_0 is Euler's constant. If $k=1$ and $\alpha=1$, C_0 is Euler's constant and $C_1=-\frac{1}{2}$. Values for L''' can be calculated for particular values of n, k and α, using (9). This process depends upon determination of an upper bound for the Bernoulli polynomial $B_{n+k}^{(k+1)}(t)$, when $0 \leqq t \leqq k+1$. This can be done from the known form of the polynomial and relations connecting the Bernoulli numbers. However, for large values of n the numerical work is excessive.

LEHIGH UNIVERSITY

ON THE UNIQUENESS OF THE SOLUTIONS OF DIFFERENTIAL EQUATIONS*

E. J. McSHANE

It is well known that the existence of solutions of systems of differential equations can be established under hypotheses not strong enough to guarantee uniqueness of the solution. The standard device for ensuring uniqueness is to assume that the functions involved satisfy a certain Lipschitz condition. Graves† showed that, for systems consisting of a single equation, this could be replaced by a certain monotoneity requirement. In this note we shall establish a uniqueness theorem which contains both of these as special cases.

We suppose that $f^1(x, y), \cdots, f^n(x, y)$ are functions defined for all x in an interval $[a, b]$ and all points $(y^1, \cdots, y^n)$ of n-space; each $f^i(x, y)$ is assumed continuous in y for each fixed x, and measurable in x for each fixed y. Under these conditions it can be shown‡ that if there exists a function $S(x)$ summable over $a \leqq x \leqq b$ such that§

$$|f(x, y)| \leqq S(x),$$

then, for each x_0 in $[a, b]$ and each point y_0, there is an absolutely continuous function $y(x) \equiv (y^1(x), \cdots, y^n(x))$ such that $y(x_0) = y_0$, and‖

$$\dot{y}^i(x) = f^i(x, y(x)), \qquad a \leqq x \leqq b, \tag{1}$$

for almost all x. However, this solution may not be unique.¶

We therefore establish the following theorem:

THEOREM I. *Let the functions $f^i(x, y)$ be defined for all (x, y) with $a \leqq x \leqq b$. Let there exist a function $M(x)$ summable over $[a, b]$ such that for all x in $[a, b]$, all y and all η, the inequality*

$$\{f^i(x, y + \eta) - f^i(x, y)\}\eta^i \leqq M(x)|\eta|^2 \tag{2}$$

holds. Then, if $y_1(x)$ and $y_2(x)$ are absolutely continuous functions satisfying the differential equations (1) *for almost all x, and for some x_0 in*

---

* Presented to the Society, December 30, 1938.

† L. M. Graves, *The existence of an extremum in problems of Mayer*, Transactions of this Society, vol. 39 (1936), pp. 456–471; in particular, p. 459.

‡ Carathéodory, *Vorlesungen über reelle Funktionen*, p. 672.

§ The symbol $|v|$ denotes the length of the vector v; thus $|f| = (f^i f^i)^{1/2}$.

‖ The symbol $\dot{y}^i$ denotes the derivative $y^{i\prime}(x)$, where that derivative exists and is finite; elsewhere it has the value 0.

¶ Carathéodory, op. cit., p. 675.

$[a, b]$ the equation $y_1(x_0) = y_2(x_0)$ holds, these functions are identical for $x_0 \leqq x \leqq b$.

Define $\eta^i(x) = y_2^i(x) - y_1^i(x)$; then $\eta(x_0) = 0$. If our theorem is false, there is a number k such that $x_0 < k \leqq b$ and $|\eta(k)| > 0$. Let h be the greatest value of x less than k for which $|\eta(x)| = 0$; such a number surely exists, since $|\eta(x)|$ is continuous and $|\eta(x_0)| = 0$. We now have

$$|\eta(h)| = 0; \qquad |\eta(x)| > 0, \qquad h < x \leqq k. \tag{3}$$

Since the functions y_1 and y_2 satisfy the equations (1), for almost all x in $[a, b]$ the equations

$$\dot{\eta}^i(x) = f^i(x, y_2) - f^i(x, y_1) = f^i(x, y_1 + \eta) - f^i(x, y_1) \tag{4}$$

hold; whence, by (2),

$$\dot{\eta}^i\eta^i \leqq \{f^i(x, y_1 + \eta) - f^i(x, y_1)\}\eta^i \leqq M(x)|\eta|^2. \tag{5}$$

If we define $\lambda(x) = \log |\eta(x)| = (\log \eta^i\eta^i)/2$, this can be written (because of (3)) in the form

$$\dot{\lambda}(x) \leqq M(x), \qquad h < x \leqq k. \tag{6}$$

On each interval $[\xi, k]$ with $h < \xi < k$, the function $|\eta(x)|$ is absolutely continuous and bounded from zero. Hence $\lambda(x)$ is also absolutely continuous on $[\xi, k]$, and from (6) we obtain by integration

$$\lambda(k) - \lambda(\xi) \leqq \int_\xi^k M(x)dx \leqq \int_a^b |M(x)|\,dx, \quad h < \xi \leqq k.$$

Thus $\lambda(\xi)$ is bounded below on the interval $h < \xi \leqq k$. But this is a contradiction; for, as x approaches h from the right, $|\eta(x)|$ approaches 0, and $\lambda(x) = \log |\eta(x)|$ approaches $-\infty$. Our theorem is therefore established.

COROLLARY. *Let the functions $f^i(x, y)$ be defined for all (x, y) with $a \leqq x \leqq b$. Let $y_1(x)$ and $y_2(x)$ be absolutely continuous functions satisfying equations* (1) *and coinciding at a point x_0 of the interval $[a, b]$. Let any one of the following five conditions be satisfied.*

(i) *To each point (x_0, y_0) with $a \leqq x \leqq b$ there corresponds a positive number ϵ and a function $M(x)$, summable over an interval $[\alpha, \beta]$ having $\alpha < x_0 < \beta$, such that*

$$\{f^i(x, y + \eta) - f^i(x, y)\}\eta^i \leqq M(x)|\eta|^2, \tag{7}$$

whenever $\alpha \leqq x \leqq \beta$ and $|\eta| < \epsilon$.

(ii) *The same condition as* (i) *with* (7) *replaced by*

$$-\{f^i(x, y + \eta) - f^i(x, y)\}\eta^i \leqq M(x)|\eta|^2. \tag{8}$$

(iii) *The same as* (i) *with* (7) *replaced by*

$$|\{f^i(x, y+\eta) - f^i(x, y)\}\eta^i| \leqq M(x)|\eta|^2. \tag{9}$$

(iv)* *The same as* (i) *with* (7) *replaced by*

$$|f(x, y+\eta) - f(x, y)| \leqq M(x)|\eta|. \tag{10}$$

(v)† $n=1$, *and* $f^1(x, y)$ *is a monotonic decreasing function of* y *for each fixed* x.

Then the identity $y_1(x) \equiv y_2(x)$ *holds on the corresponding intervals:*

(i): $x_0 \leqq x \leqq b$; (iii), (iv): $a \leqq x \leqq b$;

(ii): $a \leqq x \leqq x_0$; (v): $x_0 \leqq x \leqq b$.

The proof of (i) is essentially that of Theorem 1. We need only observe that, having found the $[\alpha, \beta]$ and the ϵ belonging to the point $(h, y(h))$, we can reduce k, if necessary, so that $[h, k] \subset [\alpha, \beta]$ and $|\eta(x)| < \epsilon$ if $h \leqq x \leqq k$. Part (ii) can be established analogously. More simply, it can be obtained from (i) by the transformation $\bar{x} = -x$. Part (iii) follows by applying (i) to the interval $x_0 \leqq x \leqq b$ and (ii) to the interval $a \leqq x \leqq x_0$. If condition (iv) holds, then, by the Cauchy-Schwarz inequality,

$$\begin{aligned}|\{f^i(x, y+\eta) - f^i(x, y)\}\eta^i| &\leqq |f(x, y+\eta) - f(x, y)| \cdot |\eta| \\ &\leqq M(x)|\eta|^2;\end{aligned}$$

so (iv) is a special case of (iii). For (v), we observe that $f^1(x, y+\eta) - f^1(x, y)$ cannot have the same sign as η; so the inequality

$$\{f^1(x, y+\eta) - f^1(x, y)\}\eta \leqq M(x)|\eta|^2$$

holds with $M(x) \equiv 0$. Hence (v) is a special case of (i).

UNIVERSITY OF VIRGINIA

* Carathéodory, op. cit., pp. 673–674.

† Graves, loc. cit.

POINT SET OPERATORS AND THEIR INTERRELATIONS*

E. C. STOPHER, JR.

INTRODUCTION

The fundamental operators of this paper will be defined in terms of a postulated derived set function. This procedure, first formally suggested by F. Riesz,† has been adopted by Chittenden‡ and others.

The notation used here, which has definite advantages over the classical notation, was first suggested by Chittenden. It first appeared in print in work by Sanders§ and later in a paper by the author.‖ Capital letters A, B, and so on, will denote arbitrary sets of points in the space S, J $(=cdS)$ will denote the isolated points of the space, and 0 $(=cS)$ the null set. Operators will be denoted by small letters d, i, and so on. Thus, for example, iA represents the interior of the set A, and dkA the derived set of the kernel of A.

Such an operator which defines a set, either null or non-null, will be called a *product operator*. The number of single operators which make up a product operator will be called the *order* of the product operator. A product operator is said to be *reduced* if it is shown to be equal to another product operator of lower order or to one of the same order but which is expressed in terms of operators which precede in the list of definitions in Part I. Those for which no reduction has been found will be called *unreduced*.

In Part I is presented a table of all the second order product operators indicating the reductions. Many other reduction formulas involving higher order operators have been found, but these will be omitted in this paper.¶ In Part II the space will be assumed to be dense in itself. With the aid of this additional postulate the theorem can be proved that all product operators of a given family can be expressed in a certain canonical form.

---

* Presented to the Society, April 9, 1937.

† F. Riesz, *Stetigkeitsbegriff und abstrakte Mengenlehre*, Atti del 4 Congresso Internationale dei Matematici, Rome, 1910, vol. 2, p. 18.

‡ E. W. Chittenden, *On general topology and the relation of the properties of the class of all continuous functions to the properties of space*, Transactions of this Society, vol. 31 (1929), pp. 290–321.

§ S. T. Sanders, Jr., *Derived sets and their complements*, this Bulletin, vol. 42 (1936), pp. 577–584.

‖ E. C. Stopher, Jr., *Cyclic relations in point set theory*, this Bulletin, vol. 43 (1937), pp. 686–694.

¶ Found with proofs in the doctor's dissertation by the author, *Interrelations of a Family of Operators on Point Sets and their Canonical Representation*, State University of Iowa, 1937.

PART I. REDUCTION OF OPERATORS

In the present discussion dA is undefined and will be called the derived set of A. We postulate:

I. $d(A+B)=dA+dB$ (operator is additive).
II. $d^2A \leqq dA$ (derived sets are closed).

We make the following definitions:*

D 1. Complement. $cA=S-A$.
D 2. Extension. $eA=A+dA$.
D 3. Interior. $iA=AcdcA$.
D 4. Concentrated part. $hA=AdA$.
D 5. Isolated set. $jA=AcdA$.
D 6. Border. $bA=AdcA$.
D 7. Frontier. $fA=AdcA+cAdA$.
D 8. Kernel. $kA=\sum B \leqq A$, such that $B \leqq dB$.
D 9. Separated part. $sA=AckA$.

The following table gives all second order operators $\phi_1\phi_2A$. The reductions are indicated and the unreduced operators are placed in parentheses.

TABLE I

ϕ_1 \ ϕ_2	c	d	e	i	h	j	b	f	k	s
c	1	(cd)	(ce)	ec	(ch)	(cj)	(cb)	(cf)	(ck)	(cs)
d	(dc)	(d^2)	d	(di)	(dh)	(dj)	(db)	(df)	(dk)	dj
e	(ec)	d	e	(ei)	(eh)	(ej)	(eb)	f	dk	ej
i	ce	(id)	(ie)	i	(ih)	ji	0	(if)	ih	si
h	(hc)	d^2	d	(hi)	(h^2)	0	(hb)	df	k	sh
j	(jc)	(jd)	j	(ji)	(jh)	j	(jb)	(jf)	0	j
b	(bc)	(bd)	(be)	0	(bh)	(bj)	b	(bf)	(bk)	(bs)
f	f	bd	be	(fi)	(fh)	(fj)	eb	bf	(fk)	(fs)
k	(kc)	(kd)	kd	(ki)	k	0	(kb)	(kf)	k	0
s	(sc)	(sd)	(se)	(si)	(sh)	j	(sb)	(sf)	0	s

One will note that of the 100 second order operators, 60 are unreduced. Of the 1000 third order operators 259 are unreduced. This decreasing rate of increase of the unreduced operators suggests that

* F. Hausdorff, *Mengenlehre*, Berlin, pp. 109–129. Under the postulates given, the extension function corresponds to Hausdorff's set of α points A_α. The points of dA correspond to his β points A_β. Similarly, hA corresponds to A_k, jA to A_i, bA to A_r (border is a translation of the German word "rand"), kA to A_k, and sA to A_s.

perhaps all product operators formed from these ten fundamental ones might be expressible in terms of some canonical system. This problem has not been solved for the general case.

It is interesting to note the simplifications which occur if we take the set A to be the whole space S. The following table shows that there are only seven distinct first order operators cS, dS, iS, jS, bS, kS and sS, and nine distinct unreduced second order ones. One might expect bS $(=SdcS)$ to be null, but it has not been assumed that the derived set of the null set is null.

TABLE II

ϕ_1 \ $\phi_2 S$	$cS=0$	dS	iS	$jS=J$	$bS=d0$	kS	sS
c	S	jS	bS	dS	iS	sS	kS
d	bS	(d^2S)	dS	(djS)	bS	kS	djS
e	bS	dS	S	(ejS)	bS	kS	ejS
i	cS	(idS)	iS	jS	cS	idS	sS
h	cS	d^2S	(hiS)	cS	bS	kS	sdS
j	cS	(jdS)	jS	jS	cS	cS	jS
b	cS	djS	cS	cS	bS	(bkS)	cS
f	bS	djS	bS	djS	bS	bkS	bkS
k	cS	kS	(kiS)	cS	bS	kS	cS
s	cS	(sdS)	sS	jS	cS	cS	sS

PART II. SPACE DENSE IN ITSELF

An additional postulate will be assumed in Part II:

III. $dS=S$.

Under Postulates I, II and III, a table similar to Table I would show 49 instead of 60 unreduced second order operators. There are 159 instead of 259 unreduced third order ones. A table similar to Table II under Postulates I, II, III would show only three distinct first order operators cS, iS, bS, and no unreduced second order operators.

Consider the family $\mathcal{E}$ consisting of the following sets, their complements, and these sets with A replaced by cA:

$$d^\alpha idA, \qquad \alpha = 0 \text{ or } 1,$$

$$h^\alpha c h^\beta A, \qquad \alpha, \beta = 1, 2, 3, \cdots \text{ (finite or transfinite ordinals)},$$

$$R(\alpha, \beta, \gamma, \kappa, \lambda, \mu, \nu) = k^\alpha s^\beta e^\gamma d^\kappa \prod_{\tau=1}^{\lambda} (jd^{\alpha_\tau}) j^\mu h^\nu A,$$

where, by definition, $\theta^0 A = A$, $\prod_{\tau=1}^{0}(jd^{a_\tau})A = A$, and R is limited by the following equations:

$$\alpha, \beta, \gamma, \mu = 0 \text{ or } 1, \qquad \alpha\beta = 0, \qquad \alpha\gamma = 0,$$
$$\alpha\kappa = 0 \text{ or } 1, \qquad \gamma\kappa = 0,$$
$$\lambda = 0, 1, 2, \cdots \qquad \text{(finite ordinals)},$$
$$\kappa, a_\tau, \nu = 0, 1, 2, \cdots \qquad \text{(finite or transfinite ordinals)}.$$

To illustrate, $R(1, 0, 0, 1, 0, 1, 5) = kdjh^5A$ and $R(0, 0, 1, 0, 2, 0, 1) = ejd^{a_1}jd^{a_2}hA$.

We will designate by $\mathcal{E}_1$ the subset of $\mathcal{E}$ for which the parameters κ and a_τ are restricted to be finite ordinals. The author has proved the following theorem:

THEOREM 1. *The set ϕA, where ϕ is any finite product operator compounded from c, d, e, i, h, j, k and s, is a set of $\mathcal{E}_1$.*

The proof is one of mathematical induction. The first part of the proof consists of showing that the set A is an $\mathcal{E}_1$ set. This is readily accomplished by noting that

$$A = R(0, 0, 0, 0, 0, 0, 0),$$

that is, A is obtained as that special case of R for which the value zero has been assigned to each of the parameters α, β, and so on.

The second part of an ordinary proof by mathematical induction would consist of showing that the successor of a general set, ϕA, of the collection $\mathcal{E}_1$ is an $\mathcal{E}_1$ set. However, each of our sets has not just one but fourteen "successors" associated with it: $d\phi A$, $dc\phi A$, $e\phi A$, $ec\phi A$, $i\phi A$, $ic\phi A$, $h\phi A$, $hc\phi A$, $j\phi A$, $jc\phi A$, $k\phi A$, $kc\phi A$, $s\phi A$, and $sc\phi A$. One can omit $ec\phi A$ and $ic\phi A$ since $ec\phi A$ equals $ci\phi A$ and $ic\phi A$ equals $ce\phi A$, but it is necessary to show that each of the other twelve "successors" of each of the $\mathcal{E}_1$ sets is an $\mathcal{E}_1$ set. This proof, which can be found in detail in the author's dissertation (loc. cit.), will be omitted here because of the large number of cases which must be considered.

The results have been extended to the following:

THEOREM 2. *The set ϕA, where ϕ is any finite product operator compounded from c, d, e, i, h, j, k, s and their transfinite powers, is an $\mathcal{E}$ set.*

Since $\theta^2 A$ equals θA or A for all of the operators except d and h, it is necessary to consider the transfinite powers of only these two. It is necessary to show that each $\mathcal{E}$ set when operated upon by transfinite powers of d or of h will yield an $\mathcal{E}$ set. For the details of this proof the reader is again referred to the same dissertation.

It remains an unsolved problem whether or not generalizations of the theorems of this paper can be established which would include the border and frontier operators or which would not require that the space be dense in itself.

ASHLAND COLLEGE

A NOTE ON A PAPER BY J. A. TODD*

A. SINKOV

In a recent paper† entitled *A note on the linear fractional group*, Todd obtained an abstract definition for the group $LF(2, 2^n)$ in terms of $n+2$ generators. Apparently he gave no consideration to the question of the independence of the defining relations, for they can be considerably simplified. First, in view of the condition $RS_i = S_{i+1}R$ (which is the same as $R^iS_0R^{-i} = S_i$), the three generators U, R and S_0 are sufficient to generate the entire group. If we give a definition in terms of these three generators alone, the relations

$$S_i^2 = 1, \qquad i \neq 0, \qquad RS_i = S_{i+1}R$$

may be discarded, and any S_i $(i \neq 0)$ which appears in the remaining conditions may be replaced by its definition in terms of R and S_0. Next, the $C_{n-1,2}$ conditions $S_iS_j = S_jS_i$ can be replaced by the $n-1$ conditions

$$S_0S_i = S_iS_0, \qquad i = 1, 2, \cdots, n-1.$$

For suppose $j-i=\alpha$. Then, from $S_0S_\alpha = S_\alpha S_0$, we get

$$R^i(S_0S_\alpha)R^{-i} = R^i(S_\alpha S_0)R^{-i}, \qquad S_iS_j = S_jS_i.$$

Writing $S_0S_i = S_iS_0$ in terms of R and S_0 only

$$S_0R^iS_0R^{-i} = R^iS_0R^{-i}S_0, \qquad (S_0R^iS_0R^{-i})^2 = 1.$$

Thus, for the three generators U, R and S_0, we require only $n+5$ conditions

$$R^{2^n-1} = U^3 = (UR)^2 = (US_0)^2 = S_0^2 = 1, \ (S_0R^iS_0R^{-i})^2 = 1,$$

$$R^nS_0R^{-n+1} = S_0^{a_0}RS_0^{a_1}RS_0^{a_2} \cdots RS_0^{a_{n-1}}R^{-n+2}.$$

But even these three generators are not independent. For the relation $(UR)^2 = 1$ permits us to consider U and R as being equivalent to

* Presented to the Society, February 25, 1939.

† Journal of the London Mathematical Society, vol. 2 (1936), pp. 103–107.

two generators of periods two and three, and these two generators can be shown to give rise to the entire group.* To put it differently, S_0 is expressible in terms of U and R. The question now arises whether the above $n+5$ conditions, when expressed in terms of U and R alone, might not involve redundancies. For example, $LF(2, 2^3)$ can be completely defined by means of only four conditions on two generators of periods two and three.† However, a general study of the independence of the relations would require the expression for S_0 in terms of U and R; this the author has been unable to obtain. It most likely would be found to involve the coefficients in the equation satisfied by the primitive root ϵ.

It nevertheless seemed an interesting problem to investigate one or two further special cases in order to see whether in those particular cases the number of defining relations could not be reduced.

Consider first $LF(2, 2^4)$. On referring to the Bulletin paper, it is seen that Todd's generators would satisfy (2, 3, 15; 17). (This is the only one of the three possibilities for which $p=15$.) Hence, the operators P and Q introduced in that paper would satisfy

$$P^{15} = (QP^2)^3 = (QP^3)^2 = (Q^3P)^2 = 1.$$

By making use of the coset enumeration process devised by Todd and Coxeter,‡ the author has been able to show that it is sufficient to add only one further relation to the above set, namely, $(Q^2P^6)^2=1$, to have a complete definition of $LF(2, 2^4)$. The basic subgroup used in the procedure was $\{Q, PQP^2\}$, of order 34, so that it was necessary to enumerate only 120 cosets. By making use of the substitution

$$P^2 = \overline{Q}, \quad Q = \overline{P}^2, \quad \overline{P}^{17} = 1,$$

it follows at once§ that a second complete definition is given by

$$P^{17} = (QP^2)^3 = (QP^3)^2 = (Q^3P)^2 = (Q^4P^4)^2 = 1.$$

In view of the fact that $LF(2, 2^4)$ has only three abstract definitions in terms of two generators of periods two and three, it seemed worth while to investigate the one remaining definition. Using the enumera-

* A. Sinkov, *On generating the simple group $LF(2, 2^N)$ by two operators of periods two and three*, this Bulletin, vol. 44 (1938), p. 455. This paper will hereafter be referred to as the Bulletin paper.

† A. Sinkov, *Necessary and sufficient conditions for generating certain simple groups by two operators of periods two and three*, American Journal of Mathematics, vol. 59 (1937), p. 70.

‡ Proceedings of the Edinburgh Mathematical Society, (2), vol. 5 (1937), pp. 26–34.

§ Compare the Bulletin paper, p. 454.

tion process once again, and with the same basic subgroup, it turned out that a complete definition is given in this instance by

$$P^{17} = (QP^2)^3 = (QP^3)^2 = (Q^9P)^2 = (Q^2P^{13})^2 = 1.$$

This now permits us to state the following necessary and sufficient condition:

THEOREM. *In order for two operators S and T of periods three and two, respectively, to generate $LF(2, 2^4)$, it is necessary and sufficient that they satisfy one of the following sets of relations:*

(1) $$P^{15} = (QP^2)^3 = (QP^3)^2 = (Q^9P)^2 = (Q^2P^8)^2 = 1,$$

(2) $$P^{17} = (QP^2)^3 = (QP^3)^2 = (Q^8P)^2 = (Q^4P^4)^2 = 1,$$

(3) $$P^{17} = (QP^2)^3 = (QP^3)^2 = (Q^9P)^2 = (Q^2P^{13})^2 = 1,$$

where $P=(ST)^{-1}$, $Q=(ST)^2S$.

If we make use of the group $G^{m,n,p}$ defined by Coxeter,* the above theorem takes a much more elegant form, since the three sets of relations reduce to two:

(4) $$G^{2,15,17},\ (C^4B^8)^2 = 1,$$

(5) $$G^{2,17,17},\ (C^4B^{13})^2 = 1.$$

In the case of $LF(2, 2^5)$, the basic subgroup $\{Q, PQP^2\}$, of order 62, required the enumeration of 528 cosets. Todd's generators satisfy the relations (2, 3, 11; 31) and a complete definition was obtained in the form

$$P^{11} = (QP^2)^3 = (QP^3)^2 = (Q^{16}P)^2 = (Q^2P^8QP^{10})^2 = (Q^2P^{10}QP^8)^2 = 1.$$

As before, this leads to a second complete definition

$$P^{31} = (QP^2)^3 = (QP^3)^2 = (Q^8P)^2 = (Q^4P^2Q^5P^4)^2 = (Q^4P^4Q^5P^2)^2 = 1,$$

both of which are equivalent to the single definition

$$G^{2,11,31},\ (C^4B^8C^2B^{10})^2 = (C^4B^{10}C^2B^8)^2 = 1.$$

We see then that when n is 3, 4 or 5, $n+1$ conditions are sufficient to yield a complete definition of $LF(2, 2^n)$. This would lead to the conjecture that the $n+5$ relations deducible from Todd's definition are, in general, not all independent.

A similar sort of treatment is possible in the case of Todd's definition for $LF(2, p^n)$. Here it follows from the definition $RS_i = S_{i+2}R$ that

* *The abstract groups* $G^{m,n,p}$, Transactions of this Society, vol. 45 (1939), pp. 73–150.

$S_2, S_3, \cdots, S_{n-1}$ are expressible in terms of either R and S_0 or R and S_1 according as the subscript is even or odd. Hence U, R, S_0, and S_1 will yield a complete definition in the case $p^n \equiv -1 \pmod 4$ if they satisfy the $2n+5$ conditions

$$R^{(p^n-1)/2} = U^3 = (UR)^2 = (US_0)^2 = S_0^p = S_1^p = 1,$$

$$S_0S_i = S_iS_0; S_1S_j = S_jS_1, i = 1, 2, \cdots, n-1; j = 2, 3, \cdots, n-1,$$

$$RS_{n-2} = S_0^{a_0}S_1^{a_1} \cdots S_{n-1}^{a_{n-1}}R, \qquad RS_{n-1} = S_0^{b_0}S_1^{b_1} \cdots S_{n-1}^{b_{n-1}}R.$$

In these relations the S_i $(i>1)$ are to be replaced by their expressions in terms of R and S_0 or R and S_1. Should $p^n \equiv +1 \pmod 4$, we must add the additional relation $(S_1RU)^3 = E$.

Here again there is the possibility that U and R alone might suffice to generate the entire group. To investigate it, we study the various subgroups of $LF(2, p^n)$ to see whether any of them can be generated by two operators satisfying the conditions $(2, 3, \{p^n-1\}/2)$. If the subgroup is commutative, or dihedral, $(p^n-1)/2$ must equal 6 or 2. If it is tetrahedral, octahedral or icosahedral, the corresponding values for $(p^n-1)/2$ are 3, 4, 5. The groups of order $p^m d$ can be eliminated by the same reasoning as was used in the Bulletin paper, after we exclude the cases $(p^n-1)/2 = 2, 3$. Linear fractional groups corresponding to Galois fields of order p^k, where k divides n are obviously impossible. Hence $(p^n-1)/2$ may not exceed 6. But n must be greater than 1. Hence the only case to consider is $p^n = 9$.

We dispose of this case at once by remembering that $LF(2, 3^2)$ can not be generated by two operators of periods two and three.* Hence, excluding this one case, U and R are sufficient to generate the entire group.

The question of the independence of the $2n+5$ conditions (or the $2n+6$ as the case may be) presents the same general difficulties that were met in the preceding considerations. Here again, however, it seemed interesting to study a special case merely to see how much improvement might be obtained in particular instances. The group chosen for study was $LF(2, 3^3)$, for which Todd's generators satisfy the relations (2, 3, 13). By the enumeration process, a complete definition was obtained by adjoining only two new conditions

$$LF(2, 3^3) \equiv (2, 3, 13), \qquad (Q^2P^5QP^9)^2 = (Q^2P^9QP^5)^2 = 1,$$

so that in this particular case six of Todd's conditions are redundant.

WASHINGTON, D. C.

* G. A. Miller, this Bulletin, vol. 7 (1900–1901), p. 426.

GROUPS OF MOTIONS IN CONFORMALLY FLAT SPACES. II

JACK LEVINE

1. **Introduction.** In a previous paper with a similar title,* we have shown that all groups of motions admitted by a conformally flat metric space V_n must be subgroups of the general conformal group G_N of $N=\frac{1}{2}(n+1)(n+2)$ parameters generated by†

$$\xi^i = b^i + a_0 x^i + x^i a_j x^j - \tfrac{1}{2} a_i e_i e_j (x^j)^2 + b_j{}^i x^j, \qquad e_i = \pm 1. \tag{1}$$

In (1), the $b_j{}^i$ satisfy the relations $e_i b_j{}^i + e_j b_i{}^j = 0$, ($i, j$ not summed). Otherwise the a's and b's in (1) are arbitrary.

To define a group of motions of V_n, the ξ^i must satisfy the equations

$$\xi^k \frac{\partial h}{\partial x^k} + h \frac{\partial \xi^i}{\partial x^i} = 0, \qquad i \text{ not summed}, \tag{2}$$

and the coordinates x^i of (2) are such that $g_{ij} = e_i \delta_j{}^i h^2$. Hence in this coordinate system, the metric has the form

$$ds^2 = h^2 \sum e_i (dx^i)^2. \tag{3}$$

In this paper we shall consider the simplest subgroups of G_N, and determine the nature of the function h corresponding to each. Also we give a restatement of Theorem 2 of I, since it is not complete as given.

2. **The group G_N.** The basis of the group G_N may be taken in the form

$$P_i = p_i, \tag{4}$$

$$S_{ij} = e_i x^i p_j - e_j x^j p_i, \qquad i, j \text{ not summed}, \tag{5}$$

$$U = x^i p_i, \tag{6}$$

$$V_i = 2x^i x^j p_j - e_i e_j (x^j)^2 p_i, \tag{7}$$

where $p_i = \partial/\partial x^i$; and its commutators are‡

* *Groups of motions in conformally flat spaces*, this Bulletin, vol. 42 (1936), pp. 418–422. The results of this paper (which we refer to as I) will be assumed known.

† All small Latin indices take the values 1, 2, $\cdots$, n, with $n>2$, unless otherwise noted.

‡ S. Lie, *Theorie der Transformationsgruppen*, vol. 3, pp. 321–334.

$$(8a) \qquad (P_i, P_j) = 0,$$

$$(8b) \qquad (P_i, U) = P_i,$$

$$(8c) \qquad (P_i, S_{jk}) = e_i\delta_{ij}P_k - e_i\delta_{ik}P_j,$$

$$(8d) \qquad (P_i, V_j) = 2\delta_{ij}U - 2e_jS_{ij},$$

$$(8e) \qquad (S_{ij}, S_{kl}) = e_j\delta_{jk}S_{il} - e_j\delta_{jl}S_{ik} - e_i\delta_{ik}S_{jl} + e_i\delta_{il}S_{jk},$$

$$(8f) \qquad (S_{ij}, U) = 0,$$

$$(8g) \qquad (S_{ij}, V_k) = e_i\delta_{jk}V_i - e_j\delta_{ik}V_j,$$

$$(8h) \qquad (U, V_i) = V_i,$$

$$(8i) \qquad (V_i, V_j) = 0.$$

The four types of symbols, P_i, S_{ij}, U, V_i, will be considered singly and in various combinations to form the subgroups to be discussed.

3. **Subgroups of one type of symbol.** We consider first the subgroups with symbols*

(a) $[P_\alpha]$, (b) $[U]$, (c) $[S_{\alpha\beta}]$, (d) $[V_\alpha]$.

The notation $[P_\alpha]$ means $[P_1, P_2, \cdots, P_r]$, and similarly for other expressions of this nature. That each of (a)–(d) forms a subgroup follows from (8a), (8e), (8i).

For (a), we have from (4), $\xi_\alpha^i = \delta_\alpha^i$, and (2), written in the form

$$\xi_\alpha^k \frac{\partial h}{\partial x^k} + h \frac{\partial \xi_\alpha^i}{\partial x^i} = 0, \qquad i \text{ not summed},$$

becomes

$$(9) \qquad \frac{\partial h}{\partial x^\alpha} = 0.$$

Hence (a): $h = h(x^{r+1}, \cdots, x^n)$. In case $r = n$, h is constant, and the V_n is flat.

The finite equations of the group $[P_\alpha]$ are

$$(10) \qquad x'^i = x^i + a^\alpha\delta_\alpha^i$$

with parameters a^α. Because of the form of (10), we call this group the T_r of *translations*. However, the group of motions $[P_\alpha]$ is not a group of translations of the V_n unless h = constant,† that is, unless V_n is flat.

* Greek letters take the values 1, 2, $\cdots$, r, with $r \leqq n$.

† L. P. Eisenhart, *Continuous Groups of Transformations*, p. 212. We refer to this book as CG.

For (b), we have $\xi^i = x^i$, and (2) becomes

$$x^i \frac{\partial h}{\partial x^i} = -h. \tag{11}$$

Hence h is homogeneous of degree -1, that is,

$$h = \frac{1}{x^1}\,\phi\left(\frac{x^2}{x^1}, \cdots, \frac{x^n}{x^1}\right), \tag{b}$$

say, where ϕ is an arbitrary function of its arguments.

The finite equations of the group $[U]$ are $x'^i = ax^i$, the group of *dilations*.

For (c), we find

$$\xi^i_{\alpha\beta} = e_\alpha \delta^i_\beta x^\alpha - e_\beta \delta^i_\alpha x^\beta, \qquad \alpha \neq \beta,$$

as the vector components of the group $[S_{\alpha\beta}]$ of $\frac{1}{2}r(r-1)$ parameters. The equations (2) which must be satisfied for each $\xi^i_{\alpha\beta}$ now become

$$X_{\alpha\beta}h \equiv e_\alpha x^\alpha \frac{\partial h}{\partial x^\beta} - e_\beta x^\beta \frac{\partial h}{\partial x^\alpha} = 0, \quad \alpha, \beta \text{ not summed}. \tag{12}$$

These equations have as general solution,

$$h = h(u; x^{r+1}, \cdots, x^n), \tag{c}$$

where $u = \sum e_\alpha (x^\alpha)^2$.

In obtaining this, we use the fact that the system (12) contains $r-1$ independent equations, since

$$e_\alpha x^\alpha X_{\beta\gamma} + e_\beta x^\beta X_{\gamma\alpha} + e_\gamma x^\gamma X_{\alpha\beta} \equiv 0, \qquad \text{no summing},$$

and it is also a complete system.*

The group $[S_{\alpha\beta}]$ has the finite equations

$$x'^\alpha = a^\alpha_\beta x^\beta, \qquad x'^A = x^A, \qquad A = r+1, \cdots, n,$$

with

$$\sum e_\alpha a^\alpha_\beta a^\alpha_\gamma = e_\beta \delta_{\beta\gamma}.$$

We call this group of $\frac{1}{2}r(r-1)$ parameters, the $R_{r(r-1)/2}$ of *rotations*.†

The vector components for the group (d) are

$$\xi^i_\alpha = 2x^i x^\alpha - e_\alpha \delta^i_\alpha R,$$

---

* Goursat, *Mathematical Analysis*, vol. 2, part 2, p. 270.

† CG, p. 57, problem 12.

where $R=\sum e_i(x^i)^2$. Equations (2) reduce, for this case, to

$$2x^\alpha x^i \frac{\partial h}{\partial x^i} - e_\alpha R \frac{\partial h}{\partial x^\alpha} + 2hx^\alpha = 0. \tag{13}$$

If we put $\lambda = x^i \partial h/\partial x^i$, (13) may be written in the form

$$\frac{2(\lambda + h)}{R} = \frac{e_\alpha}{x^\alpha}\frac{\partial h}{\partial x^\alpha}, \qquad \alpha \text{ not summed.}$$

Since the left member of this equation is independent of α, we may write

$$\frac{e_\alpha}{x^\alpha}\frac{\partial h}{\partial x^\alpha} = \frac{e_\beta}{x^\beta}\frac{\partial h}{\partial x^\beta},$$

which simplifies to (12), and hence h is of the form for (c). Using this form for h in (13), we obtain on reduction,

$$(u - v)\frac{\partial h}{\partial u} + \sum x^A \frac{\partial h}{\partial x^A} = -h, \qquad A = r+1, \cdots, n, \tag{14}$$

with

$$v = \sum e_A (x^A)^2.$$

The equation (14) has as solution

$$h = \frac{1}{R}\phi\left(\frac{x^{r+1}}{R}, \cdots, \frac{x^n}{R}\right).$$

In case $r=n$, $h=a/R$, with a constant, and the V_n is flat.*

The finite equations for the group $[V_\alpha]$ are†

$$x'^i = \frac{x^i - \frac{1}{2}R\delta_\alpha^i e_\alpha a_\alpha}{1 - a_\alpha x^\alpha + \frac{1}{4}e_\alpha e_\beta a_\alpha^2 (x^\beta)^2}.$$

4. **Subgroups with two types of symbols.** We consider in this section the simplest subgroups with two types of symbols. These are:

(e) $[P_\alpha, S_{\beta\gamma}]$, (f) $[P_\alpha, U]$, (g) $[S_{\alpha\beta}, U]$,

(h) $[V_\alpha, U]$, (i) $[S_{\alpha\beta}, V_\gamma]$.

Each of these we discuss briefly.

(e). The function h has the same form as for (a) since equations (12) are satisfied identically if (9) are.

* L. P. Eisenhart, *Riemannian Geometry*, p. 85.

† Lie, loc. cit., p. 350.

(f). Using the form of h for (a) in (11), we see that h is homogeneous of degree -1 in $x^{r+1}, \cdots, x^n$, that is, we may write

$$h = \frac{1}{x^{r+1}} \phi\left(\frac{x^{r+2}}{x^{r+1}}, \cdots, \frac{x^n}{x^{r+1}}\right).$$

If $r=n$, there is no solution.

(g). If we substitute for h in (11) its value as determined from (c), we obtain

$$2u\frac{\partial h}{\partial u} + x^A \frac{\partial h}{\partial x^A} = -h, \qquad A = r+1, \cdots, n. \tag{15}$$

Hence,

$$h = \frac{1}{u^{1/2}} \phi\left(\frac{x^{r+1}}{u^{1/2}}, \cdots, \frac{x^n}{u^{1/2}}\right).$$

(h). Equations (11) and (13) show $\partial h/\partial x^\alpha = 0$, so that h is the same as in (f). If $r=n$, there is no solution.

(i). For (d), we have seen that (13) imply (11), that is, the form of h for (i) is the same as that for (d).

5. **Subgroups with three and four types of symbols.** Of the four possibilities $[P_\alpha, S_{\beta\gamma}, V_\delta]$, $[P_\alpha, S_{\beta\gamma}, U]$, $[P_\alpha, V_\beta, U]$, $[S_{\alpha\beta}, V_\gamma, U]$, only the second and fourth give subgroups:

(j) $[P_\alpha, S_{\beta\gamma}, U]$, (k) $[S_{\alpha\beta}, V_\gamma, U]$.

For (j), the $P_\alpha, S_{\beta\gamma}$ imply $h = h(x^{r+1}, \cdots, x^n)$, and then U shows h is the same form as in (f). There is no solution of $r=n$.

The form of h for (k) will be the same for (h), as follows from (i), that is, h will have the same form as for (f). If $r=n$, there is no solution.

The simplest four type symbol subgroup is

(l) $[P_\alpha, V_\beta, S_{\gamma\delta}, U]$.

It is easily seen that the solution for h is the same as for (f), and there is no solution for $r=n$.

6. **Indices in different ranges.** So far, we have considered only subgroups whose symbol indices all have the same range, $1, \cdots, r$. In this section we discuss cases (e), (i), (j), (k), and (l) with the indices for the various types of symbols in different ranges.

Case (m): $[P_i, S_{jk}]$. Let i range through $1, \cdots, r$, and j, k through any set of t indices, $s_1, s_2, \cdots, s_t$, with $s_1 < s_2 < \cdots < s_t$. Then either:

(m_1) $s_t \leqq r$, (m_2) $s_1 \leqq r, s_t > r$, (m_3) $s_1 > r$.

For case (m_1), equations (9) imply (12) with α, β in the range $s_1, s_2, \cdots, s_t$. Hence h has the same form as in (a).

In the second case, (m_2), there must be a common index in $(1, \cdots, r)$ and $(s_1, \cdots, s_t)$, say β. Then, in (8c), choose $i=j=\beta$, and $k=s_t$. This gives

$$(P_\beta, S_{\beta s'}) = e_\beta P_{s'}, \qquad s' = s_t,$$

which is not in the set $[P_\alpha]$. Hence, this case is impossible.

For case (m_3), the two sets of indices have no index in common, and we must have $t \geqq 2$. Without loss of generality, we may take the set $s_1, \cdots, s_t$ to be $r+1, r+2, \cdots, r+t$. The form of h is easily seen to be

$$h = h(v_t; x^{r+t+1}, \cdots, x^n), \quad v_t = \sum_{r+1}^{r+t} e_J (x^J)^2.$$

Case (n): $[S_{jk}, V_i]$. As in case (m), there are three possibilities, only the first and third being possible. If we let i take the range $1, \cdots, r$, then if $s_t \leqq r$, h has the same form as for (d). If $s_1 > r$, we may let j, k have the range $r+1, \cdots, r+t$. Then h must satisfy (13), and (12) with the indices in this latter range. Since (13) implies (12), we must have $h=h(u; v_t; x^{r+t+1}, \cdots, x^n)$. Using this form for h in (13), we obtain

$$(u - w)\frac{\partial h}{\partial u} + v_t \frac{\partial h}{\partial v_t} + x^B \frac{\partial h}{\partial x^B} = -h, \qquad B = r + t + 1, \cdots, n,$$

with $w = \sum e_B (x^B)^2$. This equation has as solution

$$h = \frac{1}{R - v_t}\, \phi\left(\frac{v_t}{R - v_t}; \frac{x^{r+t+1}}{R - v_t}, \cdots, \frac{x^n}{R - v_t}\right).$$

With three types of symbols, we consider first $[P_i, S_{jk}, U]$, and let $i=1, \cdots, r$. If the indices of S_{jk} are all contained in the range $1, \cdots, r$, h has the same form as for $[P_\alpha, U]$. Otherwise, we must have all j, k indices outside the range $1, \cdots, r$. Then we have: (o) $[P_\alpha, S_{JK}, U]$, and $h=h(v_t; x^B)$, using the notation of case (n). With this value of h in (11) we obtain equation (15) with u replaced by v_t. Hence,

$$h = \frac{1}{v_t^{1/2}}\, \phi\left(\frac{x^{r+t+1}}{v_t^{1/2}}, \cdots, \frac{x^n}{v_t^{1/2}}\right).$$

As the next case we consider $[V_\alpha, S_{jk}, U]$. If the j, k indices are included in $1, \cdots, r$, we get the same form for h as in $[V_\alpha, U]$. If not

we must have j, k in the range J, K, to give: (p) $[V_\alpha, S_{JK}, U]$. The symbols V_α, U imply $h=h(x^{r+1}, \cdots, x^n)$, and then the symbols S_{JK} imply $h=h(v_t; x^B)$, the same as in (o).

The other two possibilities $[P_i, S_{jk}, V_l]$, $[P_i, V_j, U]$ are easily shown to be impossible, no matter in what ranges we choose the indices of the various symbols.

For four types we have $[P_\alpha, S_{jk}, V_l, U]$. If j, k are in the J, K range, we have a contradiction from (P_α, V_l), no matter what range l has. The only other choice is j, k included in the $1, \cdots, r$ range. Then, from (P_α, V_l), we must have l in this range also. This gives

(q) $[V_\alpha, S_{\alpha'\beta'}, V_{\gamma'}, U]$, α', β', γ' range included in $1, \cdots, r$,

and h has the same form as for (f), as easily follows.

7. **Summary.** We give here a summary of the various forms for h corresponding to the subgroups considered.

(a) $[P_\alpha]$, $h = h(x^{r+1}, \cdots, x^n)$;

(b) $[U]$, $h = \dfrac{1}{x^1}\,\phi\left(\dfrac{x^2}{x^1}, \cdots, \dfrac{x^n}{x^1}\right)$;

(c) $[S_{\alpha\beta}]$, $h = h(u; x^{r+1}, \cdots, x^n)$;

(d) $[V_\alpha]$, $h = \dfrac{1}{R}\,\phi\left(\dfrac{x^{r+1}}{R}, \cdots, \dfrac{x^n}{R}\right)$;

(f) $[P_\alpha, U]$, $h = \dfrac{1}{x^{r+1}}\,\phi\left(\dfrac{x^{r+2}}{x^{r+1}}, \cdots, \dfrac{x^n}{x^{r+1}}\right)$, $r = n$, no solution;

(g) $[S_{\alpha\beta}, U]$, $h = \dfrac{1}{u^{1/2}}\,\phi\left(\dfrac{x^{r+1}}{u^{1/2}}, \cdots, \dfrac{x^n}{u^{1/2}}\right)$;

(m_3) $[P_\alpha, S_{IJ}]$, $h = h(u; x^B)$;

(n_3) $[V_\alpha, S_{IJ}]$, $h = \dfrac{1}{R - v_t}\,\phi\left(\dfrac{v_t}{R - v_t}; \dfrac{x^B}{R - v_t}\right)$;

(o_3) $[P_\alpha, S_{IJ}, U]$, $h = \dfrac{1}{v_t^{1/2}}\,\phi\left(\dfrac{x^B}{v_t^{1/2}}\right)$;

(e) $[P_\alpha, S_{\beta\gamma}]$, and ($m_1$) $[P_\alpha, S_{\beta'\gamma'}]$, h as in (a);

(i) $[S_{\alpha\beta}, V_\gamma]$, and ($n_1$) $[V_\alpha, S_{\beta'\gamma'}]$, h as in (d);

(h) $[V_\alpha, U]$, (j) $[P_\alpha, S_{\beta\gamma}, U]$, (k) $[S_{\alpha\beta}, V_\gamma, U]$,

(l) $[P_\alpha, V_\beta, S_{\gamma\delta}, U]$, ($o_1$) $[P_\alpha, S_{\beta'\gamma'}, U]$,

(p_1) $[V_\alpha, S_{\beta'\gamma'}, U]$, (q) $[V_\alpha, S_{\beta'\gamma'}, V_{\delta'}, U]$,

all have h as in (f);

(p_3) $\qquad [V_a, S_{IJ}, U], \quad h$ as in (o_3).

In the above summary we have used the following notation:

$$R = \sum e_i(x^i)^2, \qquad u = \sum e_\alpha(x^\alpha)^2, \qquad v_t = \sum e_I(x^I)^2,$$

$i=1, \cdots, n$; Greek letters have the range $1, \cdots, r$; $I, J=r+1, \cdots, r+t$; $A=r+1, \cdots, n$; primed Greek letters have a range contained within $1, \cdots, r$; $B=r+t+1, \cdots, n$.

8. **Restatement of Theorem** 2 **of** I. In the proof of this theorem, the possibility $a_0=b^i=a_i=0$ was omitted. In this case, ξ^i has the form $\xi^i=b_j^i x^j$, and the function $f(R)$ is arbitrary. The group for this case is evidently the rotation group $[S_{ij}]$ of $\frac{1}{2}n(n-1)$ parameters. It is not difficult to show that the subgroups corresponding to the two cases mentioned in the theorem are $[ce_iP_i+V_i, S_{jk}]$ for $f(R)=(\alpha R+\beta)^2$ and $[S_{ij}, U]$ for $f(R)=\alpha R$. We may thus state the corrected theorem in the form:

THEOREM. *Every metric space with quadratic form* $\sum e_i(dx^i)^2/f(R)$ *admits the rotation group* $[S_{ij}]$ *as a group of motions. The only metric spaces with this quadratic form which admit other groups of motions are spaces of constant curvature, and* f *has the form* $f(R)=(\alpha R+\beta)^2$, *and the group is* $[ce_iP_i+V_i, S_{jk}]$, *and spaces with* $f(R)=\alpha R$, *in which case the group is* $[S_{ij}, U]$.

NORTH CAROLINA STATE COLLEGE

V_m IN S_n WITH PLANAR POINTS ($m \geqq 3$)

N. COBURN

1. **Introduction.** In this paper we shall classify the m-dimensional Riemannian manifolds (V_m) which are imbedded in an n-dimensional space of constant curvature (S_n) and whose normal curvature locus consists solely of planar points ($m \geqq 3$). Under the assumption that the second fundamental tensors have principal directions, we easily prove Segre's theorem:* *V_m in S_n with axial points are V_m in S_{m+1} or have second fundamental tensor of rank one.* Our proof is not as general as Segre's since the above additional assumption is required. However, our method can be generalized to classify the V_m in S_n with planar points. This classification is accomplished by use of the ranks of any of the two second fundamental tensors, which determine the normal curvature locus, and certain of the Ricci vectors. Our principal result is: *If the rank of any of these second fundamental tensors is greater than two, then V_m in S_n with planar points are* (1) *V_m consisting of ∞^1 V_{m-1} imbedded in ∞^1 S_{m+1}; or* (2) *V_m consisting of ∞^1 V_{m-1} imbedded in ∞^1 S_m; or* (3) *V_m lying in S_{m+2}.*

2. **Notation.** Let the unit tangent vector fields of m mutually orthogonal nonisotropic congruences of V_m in S_n be denoted by

$$\overset{c}{i}{}^{\kappa} = \epsilon \overset{c}{i}{}_{\kappa}, \qquad \epsilon = \pm 1, \qquad \kappa, \lambda, \mu = 1, \cdots, n, \tag{2.1}$$

$$\overset{d}{i}{}^{\kappa} \overset{d}{i}_{\kappa} = \overset{d}{\delta}, \qquad a, b, c = 1, \cdots, m.$$

According to whether ϵ is $+1$ or -1, we say that i^{κ} is in the positive or negative quadric of directions, determined by the first fundamental tensor of S_n ($a_{\lambda\mu}$)

$$a_{\lambda\mu} \overset{c}{i}{}^{\lambda} \overset{c}{i}{}^{\mu} = \overset{c}{\epsilon}. \tag{2.2}$$

The subscript in (2.1) refers to the congruence (orthogonal index), the contravariant index κ to the S_n coordinate system, the δ to the Kronecker symbol. For the $(n-m)$ mutually orthogonal unit vectors in the local E_{n-m} which is perpendicular to the local tangent E_m of the V_m at a point P, we write

$$\overset{p}{i}{}^{\kappa} \qquad p, q, r = m+1, \cdots, n. \tag{2.3}$$

* Most of the references are to Schouten-Struik, *Einfuhrung in die Neueren Methoden der Differentialgeometrie*, vols. 1 and 2. Noordhoff, Groningen, Batavia. Hence we shall merely indicate volume and page number: vol. 2, pp. 96, 99.

We denote the first fundamental tensor of the V_m of rank m by

$$a'_{\lambda\mu} = \sum_c \epsilon \overset{c}{i}_\lambda \overset{c}{i}_\mu = \overset{c}{i}_\lambda \overset{c}{i}_\mu. \tag{2.4}$$

Hence, the connecting unit affinor of V_m with respect to S_n becomes

$$B^\kappa_\lambda = \sum_c \epsilon \overset{c}{i}{}^\kappa \overset{c}{i}_\lambda = \overset{c}{i}{}^\kappa \overset{c}{i}_\lambda, \qquad B^{\kappa\nu}_{\lambda\mu} = B^\kappa_\lambda B^\nu_\mu. \tag{2.5}$$

For the second fundamental affinors, we write

$$\underset{p}{h}_{\lambda\mu} = - B^{\alpha\beta}_{\lambda\mu} \nabla_\alpha \underset{p}{i}_\beta = \epsilon \underset{p}{h}_{\lambda\mu}, \tag{2.6}$$

where ∇_μ denotes covariant differentiation with respect to the metric of S_n. Hence the curvature affinor is

$$H^{\cdot\cdot\kappa}_{\lambda\mu} = \underset{p}{h}_{\lambda\mu} \underset{p}{i}{}^\kappa = B^{\alpha\beta}_{\lambda\mu} \nabla_\alpha B^\kappa_\beta, \tag{2.7}$$

and for the vectors entering into the Codazzi relations (Codazzi vectors) we write

$$\underset{p}{\overset{q}{v}}_\lambda = B^\mu_\lambda (\nabla_\mu \underset{p}{i}{}^\kappa) \underset{q}{i}_\kappa = - B^\mu_\lambda (\nabla_\mu \underset{p}{i}{}^\kappa) \underset{p}{i}_\lambda. \tag{2.8}$$

Then the Gauss, Codazzi, Ricci* relations for V_m in S_n can be written

$$K'_{\alpha\beta\lambda\mu} + \kappa(a'_{\alpha\lambda} a'_{\beta\mu} - a'_{\alpha\mu} a'_{\beta\lambda}) = (\underset{p}{h}_{\beta\lambda} \underset{p}{h}_{\alpha\mu} - \underset{p}{h}_{\alpha\lambda} \underset{p}{h}_{\beta\mu}), \tag{2.9}$$

$$\nabla'_{[\mu} \underset{p}{h}_{\lambda]\alpha} + \underset{q}{\overset{p}{v}}_{[\mu} \underset{q}{h}_{\lambda]\alpha} = 0, \tag{2.10}$$

$$\nabla'_{[\mu} \underset{q}{\overset{p}{v}}_{\lambda]} + \underset{r}{\overset{p}{v}}_{[\mu} \underset{q}{\overset{r}{v}}_{\lambda]} = \underset{q}{\overset{p}{h}}_{\alpha[\mu} h_{\lambda]} \cdot^\alpha. \tag{2.11}$$

where $K'_{\alpha\beta\lambda\mu}$ denotes the Riemann-Christoffel affinor of V_m, κ the curvature of S_n, and ∇'_μ denotes covariant differentiation with respect to the metric of V_m.

3. **An imbedment theorem.** Struik† has shown that the necessary and sufficient conditions that V_m in S_n lie in a totally geodesic S_{m+k} of S_n are

$$\underset{u}{\overset{p}{v}}_\lambda = 0, \qquad p, q = m+1, \cdots, n, \tag{3.1}$$

$$\underset{u}{h}_{\mu\lambda} = 0, \qquad u, v = m+k+1, \cdots, n. \tag{3.2}$$

We shall show that a weaker form of (3.1) is ample. Let us divide the orthogonal indices as follows:

* Vol. 2, p. 130.

† Vol. 2, p. 150.

$$p, q, r = m+1, \cdots, n,$$

$$x, y, z = m+1, \cdots, m+k, \tag{3.3}$$

$$u, v, w = m+k+1, \cdots, n.$$

Now we shall prove the following theorem:

THEOREM. *The necessary and sufficient conditions that V_m in S_n lie in a totally geodesic S_{m+k} of S_n are that a set of $(n-m)$ mutually orthogonal vectors exist in the normal E_{n-m} such that*

$$\overset{z}{\underset{u}{v}}{}_{\lambda} = 0, \tag{3.4}$$

$$h_{\lambda\mu} = 0. \tag{3.5}$$

Consider the equations*

$$D_a \underset{u}{i}{}^{\lambda} = B^{\mu}_{a} \nabla_{\mu} \underset{u}{i}{}^{\lambda} = -\, h_{a\cdot}{}^{\lambda} + \overset{p}{\underset{u}{v}}{}_{a} \underset{p}{i}{}^{\lambda}. \tag{3.6}$$

By transvecting with the various unit vectors, (3.6) becomes

$$\underset{a}{i}{}^{a} D_a \underset{u}{i}{}^{\lambda} = (\text{terms in } \underset{1}{i}{}^{\lambda}, \cdots, \underset{m}{i}{}^{\lambda}) + (\overset{p}{\underset{u}{v}}{}_{a} \underset{a}{i}{}^{a}) \underset{p}{i}{}^{\lambda}. \tag{3.7}$$

Furthermore†

$$\underset{a}{i}{}^{a} D_a \underset{c}{i}{}^{\lambda} = (\text{terms in } \underset{1}{i}{}^{\lambda}, \cdots, \underset{m}{i}{}^{\lambda}) + (\underset{a}{i}{}^{a} \underset{c}{i}{}^{\beta} h_{a\beta}) \overset{p}{\underset{p}{i}}{}^{\lambda}. \tag{3.8}$$

If (3.4) and (3.5) are valid, then from (3.7), (3.8), we find

$$\underset{a}{i}{}^{a} D_a \underset{[1}{i}{}^{\lambda_1} \cdots \underset{m+k]}{i}{}^{\lambda_{m+k}} = \sigma \underset{[1}{i}{}^{\lambda_1} \cdots \underset{m+k]}{i}{}^{\lambda_{m+k}}. \tag{3.9}$$

Hence this $(m+k)$-vector determines a geodesic S_{m+k} in S_n.‡ Conversely assume V_m lies in a geodesic S_{m+k} in S_n, then (3.9) is valid. Hence one easily finds the conditions (3.4), (3.5).

4. V_m in S_n **with axial points.** We shall study the V_m in S_n with axial points. For these manifolds, the curvature affinor becomes

$$H_{\mu\lambda}{}^{\cdots\nu} = \overset{n}{\underset{n}{h}}{}_{\mu\lambda} \underset{n}{i}{}^{\nu}. \tag{4.1}$$

That is

$$h_{\mu\lambda} = 0, \qquad u, v = m+1, \cdots, n-1. \tag{4.2}$$

Hence the Codazzi relations (2.10) become, for $p = u$,

$$\overset{u}{\underset{n}{v}}{}_{[\mu} \overset{n}{\underset{n}{h}}{}_{\lambda]\alpha} = 0. \tag{4.3}$$

---

* Vol. 2, p. 130, (13.61).

† Vol. 2, p. 85, (10.5α when expanded).

‡ Vol. 2, p. 285, (13.1α). A proof could be furnished by use of vol. 1, p. 72, (6.28); vol. 1, p. 99, (10.14).

If these quantities are referred to the orthogonal congruences of V_m, then

$$\overset{u}{v}_\lambda = \overset{u}{v}_a \overset{a}{i}_\lambda, \qquad a, b, c = 1, \cdots, m, \tag{4.4}$$

$$\overset{n}{h}_{\mu\lambda} = \overset{n}{h}_{ad} \overset{a}{i}_\mu \overset{d}{i}_\lambda. \tag{4.5}$$

Thus (4.3) becomes after simplification,

$$\overset{n}{h}_{ad} \overset{u}{v}_c = \overset{n}{h}_{cd} \overset{u}{v}_a, \qquad a \neq c. \tag{4.6}$$

If the following determinant equation has real roots and simple elementary divisors*

$$| \overset{n}{h}_{\mu\lambda} - \rho a'_{\mu\lambda} | = 0, \tag{4.7}$$

then m principal directions exist such that

$$\overset{n}{h}_{ad} = 0, \qquad a \neq d, \tag{4.8}$$

is valid for the congruences determining the principal directions. Referred to these congruences, (4.6) becomes

$$\overset{n}{h}_{dd} \overset{u}{v}_a = 0, \qquad a \neq d. \tag{4.9}$$

Thus, if the second fundamental tensor has a rank greater than one, we can take $d = 1, 2$ and conclude that

$$\overset{u}{v}_a = 0, \qquad u = m + 1, \cdots, n - 1. \tag{4.10}$$

Hence, from (3.4), (3.5), these V_m lie in totally geodesic S_{m+1} of S_n. If the second fundamental tensor has rank one, then from the Gauss relations (2.9), the V_m are S_n. Furthermore, Struik† has shown in this case, from the Codazzi relations (2.10), that V_m is developable. Hence we have Struik's extension of Segre's theorem:

THEOREM. *If V_m in S_n has axial points and the rank of its second fundamental tensor is greater than one, then V_m lies in a geodesic S_{m+1}. If the rank of this tensor is one, then V_m is a developable S_m.*

5. **V_m in S_n with planar points.** Consider a V_m in S_n with planar points. Then we find similarly to (4.2)

$$\underset{u}{h}_{\lambda\mu} = 0, \qquad u, v = m + 1, \cdots, n - 2. \tag{5.1}$$

The equations corresponding to (4.6) are

$$\overset{n}{h}_{ad} \overset{u}{v}_c + \overset{n-1}{h}_{ad} \underset{n-1}{v}_c = \overset{n}{h}_{cd} \overset{u}{v}_a + \overset{n-1}{h}_{cd} \underset{n-1}{v}_a. \tag{5.2}$$

* L. P. Eisenhart, *Riemannian Geometry*, Princeton, 1926, p. 110.

† Vol. 2, p. 150.

This system can be divided into two distinct types of equations: (1) a, d, c all differ; (2) a differs from c, a coincides with d. Using the congruences for which (4.8) is valid, we write (5.2) for the case where a, d, c all differ

$$\overset{n-1}{h}_{ad}\ \underset{n-1}{\overset{u}{v}}_{c} = \overset{n-1}{h}_{cd}\ \underset{n-1}{\overset{u}{v}}_{a}, \qquad u \neq n-1, n;\ a, d, c \text{ unequal}. \tag{5.3}$$

We drop the super- and sub-indices $(u, n-1)$ while we analyze (5.3). To facilitate our work, we divide the indices into three types

$$\begin{aligned} &a, b, c = 1, 2, \cdots, m, \\ &p, q, r = 1, 2, \cdots, k, \\ &x, y, z = k+1, \cdots, m. \end{aligned} \tag{5.4}$$

The solutions of (5.3) fall into classes. In the first case, none of v_a nor h_{cd} are zero. Hence, replacing c in (5.3) by b (hence $m \geqq 3$) and dividing the resulting equation into the original one, we find

$$v_c/v_b = h_{cd}/h_{bd}, \qquad c \neq d,\ b \neq d. \tag{5.5}$$

In the second case, some of components v_p are zero. Let us assume that there are k of these (see (5.4)). Then from (5.3), we find

$$h_{pc} = 0, \qquad p \neq c, \tag{5.6}$$

unless all v_c vanish. The question arises as to whether (h_{xy}) can vanish. An easy calculation of (5.3) shows that this implies either that v_x or v_y vanishes, unless all h_{cd} vanish. Hence, no other h_{cd} than those of (5.6) can vanish unless all h_{cd} vanish. Therefore, the general solution of (5.3) is

$$v_1 = v_2 = \cdots = v_k = h_{1c} = h_{2c} = \cdots = h_{kd} = 0, \tag{5.7}$$

and (5.5) is valid for the remaining v_x, h_{xy} unless (1) all v_c vanish or (2) all h_{cd} vanish. These last two are distinct solutions of (5.3).

The number k which denotes the number of zero components of the above Ricci vector is independent of the superscript u. For if k is associated with u and k' with u' and $k < k'$, then from (5.7) $h_{xy} = 0$, $x \neq y$, $x, y > k$. But, by the above discussion, this is impossible. Hence the rank of this vector, which is k, is independent of u.

By studying the equations (5.5) and using the symmetry property of h_{cd}, we find the following solutions for case one:

$$\underset{n-1}{\overset{u}{v}}_{c} = \overset{uw}{\theta}\ \underset{n-1}{\overset{w}{v}}_{c}, \qquad u, w \neq n-1, n, \tag{5.8}$$

$$\overset{n-1}{h}_{cd} = \overset{u}{\theta}\ \underset{n-1}{\overset{u}{v}}_{c}\ \underset{n-1}{\overset{u}{v}}_{d}, \qquad c \neq d. \tag{5.9}$$

These equations are valid in case two. Hence they constitute the solution of (5.3) excepting the possibilities that v_c are all zero or h_{cd} (both of index $n-1$) are all zero for c not d. By treating the quantities of index n, we reach similar conclusions for their components formed with respect to the congruences of principal directions of the second fundamental tensor of index $(n-1)$.

If we write (5.2) for the case where a differs from c, a coincides with d, and h_{cd} (index n) vanishes for $c \neq d$, but h_{cd} and v_c (index $n-1$) do not all vanish, then from (5.9) we find

$$\overset{n}{h}_{dd}\overset{u}{v}_c + \overset{n-1}{h}_{dd}\overset{u}{v}_c = \theta \overset{u}{v}_e \overset{u}{v}_d \overset{u}{v}_d, \qquad d \neq c;\ u \neq n-1, n. \tag{5.10}$$

The solutions of (5.10) can be divided into two types: (1) the rank of $h_{\lambda\mu}$ (index n) is greater than two; (2) the rank of this tensor is less than or equal to two. In neither case may the vector v_λ (indices u, $n-1$) be a zero vector, or the two second fundamental tensors possess the same principal directions. Hence if we add to the above cases the possibilities that (3) the vector v_λ (indices u, $n-1$) is a zero vector, and (4) the two second fundamental tensors have the same principal directions, then we have listed the four types of solutions of (5.2). Evidently, the orthogonal index n may be replaced by $(n-1)$ in each of the above cases.

In this paper we shall not consider solutions of type (2). Furthermore, by use of (5.2), we find that if $h_{\lambda\mu}$ (index n) is of rank greater than two, then solutions of type (4) lead to axial points. Hence, we shall study solutions of types (1) and (3).

6. **Solutions of type** (1). If the rank of $h_{\lambda\mu}$ (index n) is greater than two, but its principal directions do not coincide with those of $h_{\lambda\mu}$ and v_λ (both $h_{\lambda\mu}$ and v_λ of index $n-1$) is not a zero vector, then the only solutions of (5.10) are

$$\overset{u}{v}_c = \beta \overset{u}{v}_c, \qquad u \neq n-1, n, \tag{6.1}$$

$$\overset{n-1}{h}_{dd} = \theta \overset{u}{v}_d \overset{u}{v}_d - \beta \overset{un}{h}_{dd}. \tag{6.2}$$

Combining this result with equations (5.8) and (5.9), we have the tensor equations

$$\overset{u}{v}_\lambda = \overset{uw}{\theta}\overset{w}{v}_\lambda, \tag{6.3}$$

$$\overset{u}{v}_\lambda = \beta \overset{u}{v}_\lambda, \tag{6.4}$$

$$\overset{n-1}{h}_{\lambda\mu} = \theta \overset{u}{v}_\lambda \overset{u}{v}_\mu - \beta \overset{uu}{h}_{\lambda\mu}. \tag{6.5}$$

From the discussion of the solutions (case 1) of (5.3), it follows that the θ terms cannot vanish.

From the Ricci relations (2.11), we find

$$\nabla'_{[\mu} \overset{u}{v}_{\lambda]} + \overset{s}{v}_{[\mu} \overset{u}{v}_{\lambda]} = 0. \tag{6.6}$$

Let us denote the $(m-1)$ congruences of V_m which are orthogonal to v_λ (index n) by

$$\overset{'g}{i}{}^{\lambda}, \qquad {}'g, {}'h = {}'2, {}'3, \cdots, {}'m. \tag{6.7}$$

By transvecting, (6.6) becomes, in virtue of (6.3), (6.4),

$$\overset{'g}{i}{}^{\lambda} \overset{'h}{i}{}^{\mu} \nabla'_{[\mu} \overset{u}{v}_{\lambda]} = 0. \tag{6.8}$$

Hence the congruence v^λ (index n) is V_{m-1} normal.

Let us denote the quantities of this V_{m-1} by barred letters. Then we choose the following normals to V_{m-1} in S_n:

$$\begin{aligned} &\overset{1}{i}{}^{\lambda} = \rho(\beta \overset{2}{i}{}^{\lambda} - \overset{n}{i}{}^{\lambda}), \qquad \rho = (1+\beta^2)^{-1/2}; \qquad \overset{n+1}{i}{}^{\lambda} = \rho(\overset{n-1}{i}{}^{\lambda} + \beta \overset{n}{i}{}^{\lambda}); \\ &\overset{n-1}{i}{}^{\lambda} = \overset{'1}{i}{}^{\lambda} = \gamma \overset{u}{v}{}^{\lambda}, \qquad \gamma = (\epsilon a_{\lambda\mu} \overset{u}{v}{}^{\lambda} \overset{u}{v}{}^{\mu})^{-1/2}; \qquad \overset{u}{i}{}^{\lambda} = \overset{u}{i}{}^{\lambda}. \end{aligned} \tag{6.9}$$

$U = m+1, \cdots, n-2$. Furthermore, let $\overline{B}_\mu^\lambda$ denote the connecting affinor of V_{m-1} in S_n

$$\overline{B}_\mu^\lambda = \overset{'g}{i}{}^{\lambda} \overset{'g}{i}_{\mu}; \tag{6.10}$$

then the second fundamental tensors of V_{m-1} in S_n are

$$\overset{p}{\bar{h}}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \nabla_\alpha \overset{p}{i}_{\lambda}, \qquad p = m+1, \cdots, n+1. \tag{6.11}$$

From (6.9), we find

$$\begin{aligned} &\overset{n}{\bar{h}}_{\lambda\mu} = \rho \overline{B}^{\alpha\beta}_{\lambda\mu} (\beta \overset{n-1}{h}_{\alpha\beta} - \overset{n}{h}_{\alpha\beta}), \\ &\overset{n+1}{\bar{h}}_{\lambda\mu} = \rho \overline{B}^{\alpha\beta}_{\lambda\mu} (\overset{n-1}{h}_{\alpha\beta} + \beta \overset{n}{h}_{\alpha\beta}), \\ &\overset{n-1}{\bar{h}}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \nabla_\alpha \overset{1}{i}_{\beta}, \\ &\overset{u}{\bar{h}}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \overset{u}{h}_{\alpha\beta} = 0, \qquad u = m+1, \cdots, n-2. \end{aligned} \tag{6.12}$$

From (6.5), we see that the V_{m-1} component of

$$\overset{n-1}{h}_{\alpha\beta} + \beta \overset{n}{h}_{\alpha\beta}$$

is zero; hence

$$\overset{n+1}{h}_{\lambda\mu} = 0, \qquad \overset{n}{h}_{\lambda\mu} = -\rho\overline{B}^{\alpha\beta}_{\lambda\mu}(\beta^2+1)\overset{n}{h}_{\alpha\beta}. \tag{6.13}$$

Thus these V_{m-1} have planar points. For their Ricci vectors which depend on $\overset{u}{\imath}^\lambda$ (index $n-1$) we find

$$\underset{n-1}{\overset{u}{v}}_\lambda = -\overline{B}^\mu_\lambda(\nabla_\mu \overset{u_\kappa}{i})\,\underset{n-1}{i}_\kappa, \qquad u = m+1, \cdots, n-1;\ u \neq n, \tag{6.14}$$

$$\underset{n-1}{\overset{n+1}{v}}_\lambda = -\overline{B}^\mu_\lambda(\nabla_\mu \overset{m+1}{\imath^\kappa})\,\underset{n-1}{i}_\kappa. \tag{6.15}$$

From the definition of this $\overset{u}{\imath}^\lambda$ and (6.14), it follows that the vectors in (6.14) are combinations of the V_m components of $\nabla_\mu i_\lambda$ (i_λ of index u). However, these components are zero. Hence the vectors in (6.14) vanish. By expanding (6.15), we find from (6.9)

$$\underset{n-1}{\overset{n+1}{v}}_\lambda = -\rho\overline{B}^\mu_\lambda \underset{1}{i}^\alpha(\nabla_\mu \underset{n-1}{i}_\alpha + \beta\nabla_\mu \underset{n}{i}_\alpha) = -\rho\overline{B}^\mu_\lambda \underset{1}{i}^\alpha(\underset{n-1}{h}_{\mu\alpha} + \beta\underset{n}{h}_{\mu\alpha}). \tag{6.16}$$

But this vanishes in virtue of (6.5). Hence we conclude

$$\underset{n-1}{\overset{u}{v}}_\lambda = 0, \qquad u = m+1, \cdots, n+1;\ u \neq n. \tag{6.17}$$

These ∞^1 V_{m-1} have planar points which are of type (3).

7. **Solutions of type** (3). Solutions of type (3) are characterized by the equations

$$\underset{n-1}{\overset{u}{v}}_\lambda = 0, \qquad u = m+1, \cdots, n-2;\ u \neq n-1, n. \tag{7.1}$$

Writing (5.2) for $a=d$, $a\neq c$, we find

$$\overset{n}{h}_{dd}\overset{u}{v}_c = 0, \qquad d \neq c. \tag{7.2}$$

These equations coincide with (4.9). As in that section, we make the following conclusions:

(7.3) If $h_{\lambda\mu}$ (index n) is of rank greater than one, then $v_\lambda = 0$ (index n). From (3.4), (3.5), these V_m lie in S_{m+2}.

(7.4) If $h_{\lambda\mu}$ (index n) is of rank one, then, say, $h_{11}\neq 0$, $v_1 \neq 0$ (index n).

Since the rank of the $h_{\lambda\mu}$ (index n) of §6 is greater than one, we have, from (7.3), the theorem:

THEOREM. *The V_m in S_n of type* (1) *consist of $\infty^1 V_{m-1}$ imbedded in $\infty^1 S_{m+1}$.*

We now study (7.4). From the equations

$$\overset{n}{h}_{\lambda\mu} = \overset{n}{h}_{11}\underset{1}{i}_\lambda \underset{1}{i}_\mu, \qquad \underset{n}{\overset{u}{v}}_\lambda = \overset{u}{\alpha}\underset{1}{i}_\lambda, \tag{7.5}$$

$$\overset{n-1}{h}_{\lambda\mu} = h_{ab}\overset{a}{i}_\lambda \overset{b}{i}_\mu, \qquad \underset{n-1}{\overset{u}{v}}_\lambda = 0, \tag{7.6}$$

and the Ricci equation

$$\nabla'_{[\mu} \overset{u}{v}_{n-1}{}_{\lambda]} + \overset{s}{v}_{n-1}{}_{[\mu} \overset{u}{v}_{s}{}_{\lambda]} = 0, \tag{7.7}$$

we find

$$\overset{u}{v}_{n-1}{}_{[\mu} \overset{u}{v}_{n}{}_{\lambda]} = 0. \tag{7.8}$$

Hence

$$\overset{n}{v}_{n-1}{}_{\lambda} = \phi \overset{}{i}_{1}{}_{\lambda}. \tag{7.9}$$

Now, using the Ricci relation,

$$\nabla'_{[\mu} \overset{n}{v}_{n-1}{}_{\lambda]} = h_{\alpha[\mu} \overset{n}{h}_{\lambda]}{}^{\alpha}, \tag{7.10}$$

and (7.9), we find

$$\overset{}{i}_{a} \overset{}{i}^{\mu} \nabla'_{[\mu} \overset{}{i}_{1}{}_{\lambda]} = 0, \qquad a, b = 2, 3, \cdots, m. \tag{7.11}$$

Hence the congruences which are orthogonal to i^{λ} (index 1) build V_{m-1}. Finally, we analyze the Codazzi relations

$$\nabla'_{[\mu} \overset{n}{h}_{\lambda]\alpha} = \overset{n}{v}_{n-1}{}_{[\mu} \overset{n-1}{h}_{\lambda]\alpha}. \tag{7.12}$$

By use of (7.5), (7.6), (7.9), these equations become

$$\overset{}{i}_{1}{}_{\alpha} \overset{}{i}_{1}{}_{[\lambda} \nabla'_{\mu]} \overset{n}{h}_{11} + \overset{n}{h}_{11} \overset{}{i}_{1}{}_{[\lambda} \nabla'_{\mu]} \overset{}{i}_{1}{}_{\alpha} + \overset{n}{h}_{11} \overset{}{i}_{1}{}_{\alpha} \nabla'_{[\mu} \overset{}{i}_{1}{}_{\lambda]} = \phi \overset{n-1}{h}_{\alpha d} \overset{a}{i}_{1}{}_{[\lambda} \overset{d}{i}_{\mu]} \overset{}{i}_{1}{}_{\alpha}. \tag{7.13}$$

By transvecting, we obtain

$$\overset{n}{h}_{11} \overset{}{i}_{b}{}^{\alpha} \overset{}{i}_{c}{}^{\mu} \nabla'_{\mu} \overset{}{i}_{1}{}_{\alpha} = \phi \overset{n-1}{h}_{cb}, \qquad c, b \neq 1. \tag{7.14}$$

Using the notation of §6 for the quantities of V_{m-1}, the equations (7.14) become

$$\overset{n-1}{h}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \nabla_{\alpha} \overset{}{i}_{\beta} = \phi (\overset{n}{h}_{11})^{-1} \overline{B}^{\alpha\beta}_{\lambda\mu} \overset{n-1}{h}_{\alpha\beta}. \tag{7.15}$$

Furthermore, if we define

$$\overset{n}{h}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \nabla_{\alpha} \overset{}{i}_{\beta}, \qquad \overset{n+1}{h}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \nabla_{\alpha} \overset{}{i}_{\beta}, \tag{7.16}$$

then from (7.5)

$$\overset{n}{h}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \overset{n}{h}_{\alpha\beta} = 0, \tag{7.17}$$

$$\overset{n+1}{h}_{\lambda\mu} = \overline{B}^{\alpha\beta}_{\lambda\mu} \overset{n-1}{h}_{\alpha\beta}. \tag{7.18}$$

From (7.15), (7.17), (7.18), we conclude that these ∞^1 V_{m-1} contain axial points. Hence the theorem:

THEOREM. *The V_m in S_n of type* (3) *are either V_m in S_{m+2} or contain ∞^1 V_{m-1} with axial points.*

If the rank of $h_{\lambda\mu}$ (index $n-1$) is greater than two, then the rank of $\overline{B}^{\alpha\beta}_{\lambda\mu}h_{\alpha\beta}$ (h of index $n-1$) is greater than one. In this case, from Segre's theorem, the ∞^1 V_{m-1} lie in ∞^1 S_m. If the rank of $h_{\lambda\mu}$ (index $n-1$) is two and its nonzero *domain* does not* contain the nonzero domain of $h_{\lambda\mu}$ (index n), then the same result is valid; if it does contain the nonzero domain of this $h_{\lambda\mu}$, then $\overline{B}^{\alpha\beta}_{\lambda\mu}h_{\alpha\beta}$ (index $n-1$) is of rank one. In this last case the ∞^1 V_{m-1} are ∞^1 developable S_{m-1}.

From §§6 and 7, we have the theorem:

THEOREM. *If the rank of any of the two second fundamental tensors is greater than two, then V_m in S_n with planar points are* (1) *V_m consisting of ∞^1 V_{m-1} imbedded in ∞^1 S_{m+1}, or* (2) *V_m consisting of ∞^1 V_{m-1} imbedded in ∞^1 S_m, or* (3) *V_m lying in S_{m+2}.*

UNIVERSITY OF TEXAS

ON TRANSITIVE GROUPS THAT CONTAIN CERTAIN TRANSITIVE SUBGROUPS†

W. A. MANNING

If a simply transitive permutation group G of compound degree n contains a regular abelian subgroup H of order n, and if at least one Sylow subgroup of H is cyclic, G is imprimitive. The proof of this important theorem, due to Wielandt,‡ is remarkable for its brevity. But familiarity with certain preliminary theorems of Schur's§ is assumed. Unfortunately these theorems, as presented by Schur, do not appear to be as elementary as they really are. It seems, therefore, worth while to offer a complete proof of Wielandt's theorem that is elementary throughout, free from the theories of rings and representations, and based on the fundamental concept of the double coset, introduced by Cauchy‖ in 1846. Some generalizations, too, can readily be made.

* Vol. 1, p. 19; German "Gebiet."

† Presented to the Society, December 29, 1938, under the title *A note on transitive groups with regular subgroups of the same degree.*

‡ H. Wielandt, Mathematische Zeitschrift, vol. 40 (1935), p. 582.

§ I. Schur, Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1933, p. 598.

‖ A. L. Cauchy, Comptes Rendus, vol. 22 (1846), p. 630.

In §4 of this paper it will be shown how, when a transitive group is given, the set of all linear homogeneous substitutions on the same n variables which are commutative with every permutation of the group can be easily and directly obtained.

1. **Double cosets and primary subsets.** The n letters permuted by the transitive group G of order ng are $a, b_1, \cdots$, and G_a is the subgroup of G that fixes the letter a. G_a is to be regarded as an intransitive group of degree n with r transitive constituents, counting those on one letter. If $ng > n$, and $r = 2$, G is doubly transitive.

If s is a permutation of G, the double coset $G_a s G_a$ is a collection* of g^2 elements, each occurring g/m times if m is the number of letters in the transitive constituent of G_a to which the letter x belongs in case s replaces a by x. If s is the identical permutation e, $G_a e G_a = G_a G_a$, which can be written G_a^2 or gG_a, a collection in which each element of the set G_a is repeated g times. The inverse double coset $(G_a s G_a)^{-1} = G_a s^{-1} G_a$ has the same number mg of distinct permutations as $G_a s G_a$, so that, if $s = (yax \cdots) \cdots$, y is one of the letters of a transitive constituent of G_a, also of degree m. This is the "pairing" of transitive constituents discovered by Burnside.† The double coset $G_a s G_a$ is unchanged if s is replaced by any other permutation of $G_a s G_a$.

Let the crosscut of two finite collections K and L be indicated by the symbol $K \cap L$, and be defined by saying that if an element is repeated k times in K and l times in L, it occurs $(k+l-|k-l|)/2$ times in $K \cap L$.

Let H be a transitive subgroup of degree n and order nh of the transitive group G. The subgroup of H that fixes a is H_a, and is of order h. The crosscut $H \cap (G_a s G_a)$ will, if s is a permutation of G, be called a primary subset of H with respect to G_a. The subset H_a of H is a primary subset of H with respect to G_a, even when $h = 1$. If $s = s_1 = (ab_1 \cdots) \cdots$, where b_1 is one of the m letters $b_1, b_2, \cdots, b_m$, of a transitive constituent of G_a, this crosscut consists of mh distinct permutations of H:

$$K: \qquad s_1 = (ab_1 \cdots) \cdots, \; s_2 = (ab_2 \cdots) \cdots, \; \cdots,$$
$$s_m = (ab_m \cdots) \cdots; \; \cdots.$$

If a collection consists of these mh permutations of a primary subset,

* A collection becomes a set if we disregard the repetitions of an element. Cf. Kestleman, *Theories of Integration*, 1938.

† W. Burnside, Proceedings of the London Mathematical Society, vol. 33 (1901), p. 162.

each repeated n_1 times, it is written n_1K. A chief collection of H with respect to G_a is a collection of primary subsets of H with respect to G_a, and if $K_1, K_2, \cdots$ are the primary subsets involved in the chief collection K, we may use the notation

$$K = n_1K_1 + n_2K_2 + \cdots + n_qK_q,$$

where $n_1, n_2, \cdots$ are natural numbers. A primary subset as here defined is, when H is regular, equivalent to the "primären Komplex" of Schur, Wielandt and Kochendörffer, but as there is confusion in the use of the word "Hauptkomplex," I prefer to avoid the word "complex" altogether.

It results immediately from the definition that H has exactly r primary subsets, with respect to G_a corresponding to the r transitive constituents of G_a, and any given permutation of H is in one and only one primary subset (with respect to G_a). If K is a primary subset of H with respect to G_a, it is a primary subset of H with respect to every subgroup G_x that fixes one letter x of G if K is invariant under H.* In particular, the phrase "with respect to G_a" can be omitted when H is abelian. Two collections K and L of permutations of G are said to be permutable if $KL = LK$, exact account being taken of the number of times a permutation recurs. With this convention as to the product of two collections we can state a well known theorem as follows:

Two permutation groups U and V, of orders u and v respectively, with w permutations in common, generate a group of order uv/w if and only if $UV = VU$. Then H and G_a are permutable groups.

2. **Criteria for primitivity.** Throughout this section we shall retain the preceding notation: H is any transitive subgroup of degree n and

* The necessary and sufficient condition that a primary subset K of H with respect to G_a be a primary subset of H with respect to every subgroup G_x is that all the transforms of K under H are also primary subsets of H with respect to G_a.

This theorem can be proved as follows. Let $K = H \cap G_xAG_x$ for every letter x of G. Let $t = (xa \cdots) \cdots$ be a permutation of H. Then $t^{-1}Kt = H \cap G_at^{-1}AtG_a$, because $t^{-1}G_xt = G_a$. Note that for $(xa \cdots) \cdots$ we can use any one of the permutations tH_a, because the permutations of H_a transform every primary subset of H with respect to G_a into itself, and $t^{-1}Kt$ is seen to be a primary subset of H with respect to G_a by definition. Now if $K = H \cap G_xAG_x$ for every x, the permutations of a certain number of cosets $tH_a, \cdots$ transform K into K, thus forming a subgroup N of H, and the permutations $t_1N, t_2N, \cdots$ transform K into the primary subsets $K_1, K_2, \cdots$ with respect to G_a. Conversely, if all the transforms of K under H are primary subsets of H with respect to G_a we take n permutations $t = (yax \cdots) \cdots$ from H, y running through the n letters. Now $t^{-1}Kt = H \cap G_xt^{-1}AtG_x = H \cap G_at^{-1}AtG_a$, the last because $t^{-1}Kt$ is a primary subset of H with respect to G_a, by hypothesis. Hence $K = t(t^{-1}Kt)t^{-1} = H \cap G_yAG_y$, for every letter y of G. June 20, 1939.

order nh of a transitive group G of degree n and order ng: H_a and G_a are the subgroups fixing one letter a of H and G respectively. Our fundamental theorem is the following:

THEOREM 1. *A collection K of permutations of H is a chief collection of H with respect to G if and only if $H_aK = hK$ and $G_aK = KG_a$.*

First, let K be a primary subset of H with respect to G_a composed of the mh permutations of $H_as_1, H_as_2, \cdots, H_as_m$, where

$$s_1 = (c_1ab_1 \cdots) \cdots, \quad s_2 = (ab_2 \cdots) \cdots, \quad \cdots,$$
$$s_m = (ab_m \cdots) \cdots.$$

Evidently $H_aK = hK$. The set of mg permutations in the m cosets $G_as_1, G_as_2, \cdots, G_as_m$ consists of the distinct permutations of the double coset $G_as_1G_a$; therefore, exactly the same permutations occur in $s_1G_a, s_2G_a, \cdots, s_mG_a$. Thus $G_aK = KG_a$, when K is a primary subset. If $K = H_a$, $H_aK = hK$ and $G_aK = KG_a$. It follows that, if K is a chief collection of H with respect to G_a, $H_aK = hK$ and the collections G_a and K are permutable.

To prove the converse, let K be a collection of permutations of H which satisfies the two conditions of the theorem. Then $G_aG_aK = G_aKG_a$, and we expand this by putting for K each of its permutations (call them $s_1, s_2, \cdots$), each as often as it occurs in K, so that

$$gG_aK = n_1(G_as_1G_a) + n_2(G_as_2G_a) + \cdots.$$

The only permutations of H in G_aK are the permutations of H_aK, which by hypothesis is hK; therefore

$$ghK = n_1(H \cap G_as_1G_a) + n_2(H \cap G_as_2G_a) + \cdots,$$

that is, ghK is a chief collection. Therefore K is a chief collection with respect to G. For, if K is not a chief collection, ghK is not one.

THEOREM 2. *If K and L are two chief collections of H with respect to G_a, KL is a chief collection with respect to G_a.*

For $HKL = hKL$ and $G_aKL = KG_aL = KLG_a$.

A chief collection of H is said to be singular if its permutations generate a proper subgroup of H of order greater than h.

THEOREM 3. *The group G is imprimitive if and only if some chief collection of H is singular.*

Let G be imprimitive. Then G_a is a proper subgroup of a proper subgroup J of G. This J does not include H, but it does contain at

least one permutation of the set $H-H_a$. Let s be a permutation of $J-G_a$. Since every element of G_asG_a is in J, $H \cap G_asG_a$ is a subset of $H \cap J$, which last is a proper subgroup of H different from H_a; hence $H \cap G_asG_a$ is a singular primary subset.

Let K be a singular chief collection of H with respect to G_a. Its elements generate a proper subgroup C of H of order greater than h. Since C is of finite order, there exists a natural number N such that every permutation of C is in the chief collection $K+K^2+\cdots+K^N$; of course every element of this collection is in C, so that C is a chief collection of H with respect to G_a. By Theorem 1, H_a is a subgroup of C, and $CG_a=G_aC$, that is, C and G_a are permutable groups with H_a in common, and therefore CG_a is a proper subgroup of G, which means that G is imprimitive.

The following theorem is an extension of Wielandt's lemma:

THEOREM 4. *Let P be a proper subgroup of H but not a subgroup of H_a. If the elements of a set of left (or right) cosets of H with respect to P constitute a chief collection K of H with respect to G_a, and if $K \neq H$, G is imprimitive.*

If P alone, or if P and certain cosets with respect to P, constitute a chief collection L of H ($L \neq H$), then $K=H-L$ is a chief collection composed of cosets with respect to P. Hence it is sufficient to discuss the case in which K does not contain P. Then there are at least two primary subsets involving permutations of P (e is in one of them) and at least one in K, so that H has at least three primary subsets with respect to G_a, which means that G_a has at least three transitive constituents. Let $H_as_1, H_as_2, \cdots, H_as_k$, where $s_1=(ay_1\cdots)\cdots$, $s_2=(ay_2\cdots)\cdots, \cdots, s_k=(ay_k\cdots)\cdots$, be the permutations of the collection K. Now $H_as_1P, H_as_2P, \cdots, H_as_kP$ include (with repetitions) all the left cosets that make up K. Let $t=(y_iz\cdots)\cdots$ be a permutation of P. Since the product $s_it=(az\cdots)\cdots$ belongs to K, z is one of the letters $y_1, y_2, \cdots, y_k$, and therefore P permutes the letters of P only among themselves. But G_a also permutes these letters $y_1, y_2, \cdots, y_k$ among themselves and in consequence the group generated by P and G_a is intransitive. Hence G_a is a proper subgroup of a proper subgroup of G, and G is imprimitive.

The following theorem is due to Schur, and the proof which we shall briefly indicate is his.*

THEOREM 5. *Let G be a primitive group and let the transitive group H be abelian. If p is a prime divisor of the degree n, and if r is any permuta-*

* I. Schur, loc. cit.

tion of H other than the identity, the number of solutions of $x^p=r$ in a primary subset K of H is congruent to zero, modulo p.

Let $s_1, s_2, \cdots, s_m$ be the permutations of the primary subset K. In the chief collection K^p the number of times any element, except perhaps $s_1^p, s_2^p, \cdots, s_m^p$, occurs is, because H is abelian, a multiple of p. If from K^p we remove those primary subsets which occur kp times ($k=1, 2, \cdots$), the remaining collection, if not empty, will be a chief collection of H consisting of some or all of the permutations $s_1^p, s_2^p, \cdots, s_m^p$. Now $s_1^p, s_2^p, \cdots, s_m^p$ generate a proper subgroup of H, and therefore the chief collection remaining generates a proper subgroup J of H, which is not allowable unless $J=e$ because, by hypothesis, G is primitive. Hence after the removal of certain permutations because they occur in sets of p, only e, the identity, can remain. Thus $x^p=r$, where $r\neq e$, either has no solution in K or kp of the permutations $s_1, s_2, \cdots, s_m$ of K satisfy it.

3. **Wielandt's theorem.** We are now prepared to prove

THEOREM 6. *If a simply transitive group G not of prime degree contains a regular abelian subgroup H of the same degree, and if one of the Sylow subgroups of H is cyclic, G is imprimitive.*

Let K be a primary subset (not e) of H, and let p be the prime to which there corresponds a cyclic Sylow subgroup of H. If x_0 is an element of K such that $x_0^p=r$, and if x_0y is another solution in K, $y^p=e$. But because H has only one subgroup of order p, only p permutations of H satisfy $y^p=e$. If $P=\{s\}$ is this cyclic subgroup of order p, $x_0, x_0s, \cdots, x_0s^{p-1}$ are the only possible solutions of $x^p=r$ in K. Therefore, if K contains no permutation of order p, and if G is primitive, all the permutations of K lie in the cosets $x_0P, x_1P, \cdots$, so that by Theorem 4 the group G should be imprimitive. Since H has at least three primary subsets, we can, when $p=2$, choose K so that $K \cap P$ is empty. In the following, then, p is an odd prime. If K is a subset of P, K is singular because of the condition that p is less than n; this again would make G imprimitive.

If G is primitive, $K=XP+F$, where the set X consists of permutations $x_0, x_1, \cdots$ of K whose pth powers are in $H-e$, and no two of which have the same pth power; and the set F consists of the f distinct permutations $s^{n_1}, s^{n_2}, \cdots, s^{n_f}$ of order p. Neither X nor F is empty. Now $P^2=pP$, $PF=fP$, and in F^2 no permutation can occur more than f times. Hence (H being abelian) $K^2=pX^2P+2fXP+F^2$.

Let us say that we have chosen for our primary subset K a primary subset of H for which $f\leqq(p-1)/2$. Since K^2 is a chief collection and

$2fXP$ contains a permutation of K repeated at least $2f$ times, K^2 contains $2fK=2fXP+2fF$. Then K^2-2fK is a chief collection and is equal to $(pX^2P+F^2)-2fF$. No element of F^2 occurs more than f times, so that every element of F is an element of X^2P. If then from $(pX^2P+F^2)-2fF$ we remove primary subsets p at a time until no permutation left occurs more than $p-1$ times, we get a chief collection $F^2+(p-2f)F$, clearly singular, implying that G is imprimitive.

Thus Wielandt's theorem is proved, and the proof is Wielandt's, stripped to show its truly elementary character.

4. **Linear substitutions.** Let G be a transitive group on the n letters $x_1, x_2, \cdots, x_n$. Let G_k be its subgroup that fixes the one letter x_k. We agree as before that G_k is to be regarded as a group of degree n with r transitive constituents. Let one of the transitive constituents of G_1 be on the m letters (or variables if you like) $x_{11}, x_{12}, \cdots, x_{1m}$, where x_{1j} is of course one of the letters $x_1, x_2, \cdots, x_n$. Let t_k be a permutation of G that replaces the $m+1$ letters $x_1, x_{11}, x_{12}, \cdots, x_{1m}$ by $x_k, x_{k1}, x_{k2}, \cdots, x_{km}$, respectively. Then $t_k^{-1}G_1t_k=G_k$, and in G_k is found a transitive constituent on $x_{k1}, x_{k2}, \cdots, x_{km}$. Thus there is a one-to-one correspondence between the n subgroups $G_1, G_2, \cdots, G_n$ and the n sets of letters $x_{1j}, x_{2j}, \cdots, x_{nj}$ $(j=1, 2, \cdots, m)$. Then the linear substitution

$$
\begin{aligned}
x_1' &= x_{11} + x_{12} + \cdots + x_{1m}, \\
x_2' &= x_{21} + x_{22} + \cdots + x_{2m}, \\
&\cdots\cdots\cdots\cdots\cdots, \\
x_n' &= x_{n1} + x_{n2} + \cdots + x_{nm}
\end{aligned}
\tag{1}
$$

is transformed into itself by every permutation of G. Neither the n subgroups G_k nor the n sets of m letters need be distinct. In this way r substitutions can be set up, and if each of them is multiplied by an arbitrary constant and they are then added, the result is a substitution commutative with every permutation of G. If $m=1$, (1) is a permutation. If G is a regular group, $r=n$, and we get n distinct permutations, that is, the conjoin of G.

Let us suppose now that the substitution

$$
V: \qquad x_i' = \sum_{j=1}^{n} v_{ij}x_j, \qquad i = 1, 2, \cdots, n,
$$

is commutative with every permutation of G. Consider the first equation $x_1'=\sum_{j=1}^{n}v_{1j}x_j$ of V. When V is transformed by all the permutations of G_1, x_1' is invariant and therefore the m coefficients of the

variables $x_{11}, x_{12}, \cdots, x_{1m}$ must be equal, to v_1, say. Next, consider the transform of V by t_k (defined in the preceding paragraph): x_1' goes into x_k', and the sum $v_1(x_{11}+x_{12}+\cdots+x_{1m})$ goes into $v_1(x_{k1}+x_{k2}+\cdots+x_{km})$, so that the variables $x_{k1}, x_{k2}, \cdots, x_{km}$ in $x_k'=\sum_{j=1}^{n} v_{kj}x_j$ all have the same coefficient v_1, for $k=1, 2, \cdots, n$. Hence:

THEOREM 7. *Every linear substitution commutative with each of the permutations of a transitive group can be obtained by first setting up the substitutions* (1) *and then multiplying them by appropriate constants and adding.*

If the given transitive group G contains a regular subgroup H of the same degree n, the permutations $t_1=e, t_2, t_3, \cdots, t_n$ can be the permutations of H, and the substitution (1) will be the sum of m permutations of the conjoin of H. There is a connection between the substitution (1) and a primary subset of H.

THEOREM 8. *Let G be a transitive group of degree n and H a regular subgroup of order n. Let the primary subset K of H be transformed into itself by every permutation of H. Then the sum of the permutations of K is commutative with every permutation of G.*

This K is independent of our choice of subgroup fixing one letter. We return to the notations of §1. If $H, H_1, \cdots, H_{w-1}$ is the complete set of conjugates to which H belongs under G, it is also, because $HG_x=G$, a complete set of conjugates under G_x. Those permutations of G_x which transform H into H, transform K into K. Hence K has one and only one primary subset conjugate to it in each of the subgroups $H, H_1, \cdots, H_{w-1}$. These w primary subsets $K, K_1, \cdots, K_{w-1}$, whether distinct or not, constitute a conjugate set under G. If in the permutations of K the letter x is followed by $y_1, y_2, \cdots, y_m$, it is followed by the same m letters in some order in K_1, in $K_2, \cdots$, in K_{w-1}. Therefore the sum of the permutations of K_i is equal to the sum of the permutations of K $(i=1, 2, \cdots, w-1)$.

This result applies to every primary subset of H when H is abelian. The following theorem is more general.

THEOREM 9. *Let G be a transitive group of degree n and H a regular subgroup of order n. Let G_a be the subgroup of G that fixes the one letter a, and let K be a primary subset of H with respect to G_a, composed of the m permutations*

$$t_1 = (c_1ab_1\cdots)\cdots, t_2 = (c_2ab_2\cdots)\cdots, \cdots,$$
$$t_m = (c_mab_m\cdots)\cdots.$$

Let K^* be the subset of the conjoin of H composed of the m permutations

$$t_1^* = (ac_1 \cdots) \cdots, t_2^* = (ac_2 \cdots) \cdots, \cdots,$$
$$t_m^* = (ac_m \cdots) \cdots.$$

Then the sum of the permutations of K^* is a linear substitution commutative with every permutation of G.

There exists a permutation $(a)(b_1c_1)(b_2c_2) \cdots (b_mc_m) \cdots$ of order 2 which transforms the regular group H into its conjoin H^*.

Any permutation s that transforms H into H_i, transforms H^* into the conjoin H_i^* of H_i; because if every permutation of H^* is commutative with every permutation of H, then every permutation of $s^{-1}H^*s$ is commutative with every permutation of $s^{-1}Hs$. Let N_a be the crosscut of G_a and N, the normalizer of H in G. The permutations of N_a transform each primary subset K of H with respect to G_a into itself. The w conjugate subgroups $H, H_1, \cdots, H_{w-1}$ contain w primary subsets with respect to G_a: $K, K_1, \cdots, K_{w-1}$, forming, though perhaps not all distinct, a conjugate set under G_a, as do also the corresponding subsets $K^*, K_1^*, \cdots, K_{w-1}^*$ of $H^*, H_1^*, \cdots, H_{w-1}^*$ (respectively). The permutation $t = (ax \cdots) \cdots$ of H transforms G_a into G_x, and therefore replaces the letters $c_1, c_2, \cdots, c_m$ of a transitive constituent of G_a by the letters $z_1, z_2, \cdots, z_m$ of a transitive constituent of G_x. But t transforms each permutation of K^* into itself. Hence in K^*, x is followed by $z_1, z_2, \cdots, z_m$. Now

$$G_a = N_a + N_aq_1 + N_aq_2 + \cdots + N_aq_{w-1},$$

where each permutation of N_aq_i transforms K^* into K_i^*. The permutation q_i is in HG_x, and every permutation of H is commutative with every permutation of K^*, while every permutation of G_x permutes among themselves the m letters $z_1, z_2, \cdots, z_m$ that follow x in the permutations of K^*. This being true for every letter x of G, and for every i, the sum of the permutations of K_i^* is equal to the sum of the permutations of K^*.

STANFORD UNIVERSITY

QUADRATIC AND LINEAR CONGRUENCE*

R. E. O'CONNOR, S.J.

The number of simultaneous solutions of a quadratic and a linear congruence does not seem to be discussed in the literature, yet a knowledge of the invariants necessary to specify this number should lead to an arithmetical classification of the form-pairs involved. This preliminary investigation is confined to congruences with modulus odd and prime to the g.c.d.'s of the two sets of coefficients. From the formulas obtained, a simple use of the Chinese Remainder Theorem will give the number of solutions for any such modulus which is either square-free or at least whose prime factors of power greater than the first are of a definite class. An interesting application, of a different type from the preceding, is given in §6. Special cases of this and of Theorem 1 have already been proven.†

1. **Hypotheses and definitions.** We shall be considering *the number* $N(p^m)$ *of simultaneous solutions of the congruences*

$$(1^m) \qquad f(x) = \sum_1^n a_{ij}x_ix_j \equiv r, \qquad g(x) = \sum_1^n c_ix_i \equiv s \pmod{p^m}$$

with f and g integral forms, $n \geqq 2$, *r and s integers, and p an odd prime dividing neither the g.c.d. of the coefficients of f nor that of g. Defining* $\phi(x, t) = f(x) + 2tg(x)$, *let a be the determinant of f,* μ *be the modulo p rank of a, b be the determinant of* ϕ, ν *be the modulo p rank of b, and* $k = s^2a + rb$.

With the above forms are to be associated three others—$F(x)$, $G(x)$ and $\Phi(x, t) = F(x) + 2tG(x)$—related to the above as follows. By a well known theorem‡ we can find a linear, integral transformation T of determinant unity that takes f into a form f' which is congruent (mod p^m) to a form

$$F(x) = a_1x_1^2 + \cdots + a_nx_n^2,$$

where $p \mid a_1a_2 \cdots a_\mu$, $p \mid a_{\mu+1}, a_{\mu+2}, \cdots, a_n$. The transformation T', identical with T for the variables x and taking t into itself, is also unimodular and takes $\phi(x, t)$ into the form $f'(x) + 2tG(x)$, where

$$G(x) = b_1x_1 + \cdots + b_nx_n.$$

---

* Presented to the Society, April 14, 1939.

† G. Pall and R. E. O'Connor, American Journal of Mathematics, vol. 61 (1939), pp. 491–496.

‡ Minkowski, *Gesammelte Abhandlungen*, vol. 1, p. 14.

Let A be the determinant of F and B the determinant of Φ.

If we consider now the congruences

$$(2^m) \qquad F(x) \equiv r, \qquad G(x) \equiv s \pmod{p^m},$$

the following facts are clear. The prime p divides neither the g.c.d. of the a_i's nor that of the b_i's; $A \equiv a$, $B \equiv b \pmod{p^m}$; the modulo p ranks of the matrices of F and Φ are respectively μ and ν; and the number $N(p^m)$ of solutions of (1^m) is the number of solutions of (2^m).

A *singular solution* of (1^m) is a solution ξ such that, for some integer λ (including zero), p divides each of the n integers $\sum_j (a_{ij}\xi_j) - \lambda c_i$ for $i = 1, 2, \cdots, n$. From the fact that this expression is obtained by setting $t = -\lambda$ in the partial derivative of $\phi(x, t)/2$ with respect to x_i, it is easily seen that the transformation T defines a one-to-one correspondence between the singular solutions of (1^m) and those of (2^m). Solutions of the congruences which are not singular we shall call *ordinary*.

2. **Inhomogeneous quadratic congruence.** We shall need explicit formulas for the number Q of solutions of

$$(3) \qquad \sum_1^n \alpha_{ij}x_ix_j - 2s\sum_1^n \beta_i x_i \equiv c \pmod{p}$$

in several hypotheses.

LEMMA.* *Let p be an odd prime; let s, c, α_{ij} $(=\alpha_{ji})$, and β_i, $(i, j = 1, 2, \cdots, n)$, be assigned integers; let δ be the determinant $|\alpha_{ij}|$ and δ' the integer $-\sum_1^n A_{ij}\beta_i\beta_j$, where A_{ij} is the cofactor of α_{ij} in δ.*

(4_1) If $p \nmid s\delta'$, $p \mid \delta$, then $Q = p^{n-1}$.

(4_2) If $p \nmid \delta'$, $p \mid \delta$, s, then, according as n is even or odd, $Q - p^{n-1} = p^{(n)}((-1)^{(n)}\delta' c \mid p)$ or $p^{(n)}p[c]((-1)^{(n)+1}\delta' \mid p)$.

(4_3) If $p \nmid \delta, \cdots$, then, according as n is even or odd, $Q - p^{n-1} = p^{(n)-1}p[\sigma]((-1)^{(n)}\delta \mid p)$ or $p^{(n)}((-1)^{(n)}\sigma \mid p)$.

Here (n) in the indices stands for $n/2$ or $(n-1)/2$ according as n is even or odd; σ for $\delta c - \delta' s^2$; and $p[z]$ for $p - 1 - p(z^2 \mid p)$ and hence for $p-1$ or -1 according as $p \mid z$ or $p \nmid z$. These formulas are valid for $n \geqq 1$, as is evident from the proof except in the case (4_2) with $n = 1$; but for this case (4_2) is easily verified directly.

By the process used to replace ϕ by Φ in §1, we can replace (3) by

* Formulae (4_1) and (4_2) for the case $n = 2$ are given by R. LeVavasseur (Mémoires de l'Académie des Sciences de Toulouse, (10), vol. 3 (1903), p. 39). A confusion of sign in summary on p. 46 has led to a similar confusion in Dickson, *History of the Theory of Numbers*, vol. 2, p. 327.

$$\sum_1^n \alpha_i y_i^2 - 2s\sum_1^n \gamma_i y_i \equiv c \pmod{p}, \tag{5}$$

which has the same number of solutions as (3) and where

$$\begin{aligned} &\alpha_1\alpha_2\cdots\alpha_n \equiv \delta, \quad \gamma_1^2\alpha_2\alpha_3\cdots\alpha_n + \gamma_2^2\alpha_3\alpha_4\cdots\alpha_n\alpha_1 + \cdots \\ &\qquad + \gamma_n^2\alpha_1\alpha_2\cdots\alpha_{n-1} \equiv -\delta' \pmod{p}. \end{aligned} \tag{6}$$

From (6) it is clear that the hypothesis $p\mid\delta$ and $p\nmid\delta'$ implies that p divides exactly one of the α_i's, say α_n, and that p does not divide the corresponding γ_i, that is γ_n, and consequently that $\delta'\equiv-\gamma_n^2\alpha_1\alpha_2\cdots\alpha_{n-1} \pmod{p}$. Thus in the hypothesis of (4_1) the left side of (5) contains a nonzero linear term, say $2s\gamma_n y_n$, such that the corresponding quadratic coefficient vanishes (mod p). The numbers $y_1, y_2, \cdots, y_{n-1}$ may be chosen arbitrarily and y_n is uniquely determined to satisfy (5).

In the hypothesis of (4_2), one quadratic and all linear coefficients of (5) vanish (mod p). Let $p\mid\alpha_n$. The number of solutions of (5) is then p times that of $\sum_1^{n-1}\alpha_i y_i^2\equiv c \pmod{p}$. This gives* (4_2) in view of the remark immediately following (6).

To establish (4_3) we can transform (5)—since no α_i vanishes—into $\sum_1^n \alpha_i z_i^2 \equiv c+Rs^2 \pmod{p}$ by the transformation $y_i=z_i+sa_i\gamma_i$, $(i=1, 2, \cdots, n)$, where a_i is some integer satisfying $a_i\alpha_i\equiv 1 \pmod{p}$. Here R will be $\sum a_i\gamma_i^2$, so by (6) $R\delta\equiv-\delta' \pmod{p}$. The number of solutions z is then given by (4_3).

3. $N(p)$ **for simple invariants.** The following theorem is required for the proof of the more general Theorem 2 but in certain applications is more useful than the latter.

THEOREM 1. *With the hypotheses and definitions of* §1,

(7_1) *if* $p\nmid as$, $p\mid b$, *then* $N(p)=p^{n-1}$;

(7_2) *if* $p\nmid a$, $p\mid b$, s, *then, according as* n *is even or odd,* $N(p)-p^{n-2}=p^{(n)-1}p[r]((-1)^{(n)}a\mid p)$ *or* $p^{(n)}((-1)^{(n)}ar\mid p)$;

(7_3) *if* $p\nmid b$, $\cdots$ *then, according as* n *is even or odd,* $N(p)-p^{n-2}=p^{(n)-1}((-1)^{(n)}k\mid p)$ *or* $p^{(n)-1}p[k]((-1)^{(n)+1}b\mid p)$.

The superior symbol (n) *and the symbol* $p[\ ]$ *are defined in the lemma.*

* Here and later we make use of the following theorem: *The number of solutions of the quadratic congruence* $u_1x_1^2+u_2x_2^2+\cdots+u_nx_n^2\equiv h \pmod{p}$, *where* p *is an odd prime,* $p\nmid u=u_1u_2\cdots u_n$, h *an arbitrary integer and* $n\geq 1$, *is given by the formula,* $p^{n-1}+p^{(n-s)/2}p[h]((-1)^{n/2}u\mid p)$ *or* $p^{n-1}+p^{(n-1)/2}((-1)^{(n-1)/2}uh\mid p)$ *according as* n *is even or odd.* Here we have written $p[h]$ for $p-1-p(h^2\mid p)$ and $(s\mid p)$ is a Legendre symbol. These formulae were first given by C. Jordan in Comptes Rendus de l'Académie des Sciences, Paris, vol. 62 (1866), p. 687.

Replacing (1^1) by (2^1), we may suppose $p \nmid b_\tau$. Solving the linear congruence of (2^1) for x_τ and substituting this in the quadratic, we obtain, on multiplying the resulting eliminant by b_τ^2,

$$(8)\qquad \sum{}'(b_\tau^2 a_i + b_i^2 a_\tau)x_i^2 + a_\tau \sum_{i\neq j}{}' b_i b_j x_i x_j - 2sa_\tau \sum{}' b_i x_i \equiv b_\tau^2 r - a_\tau s^2 \pmod p,$$

where $\sum'$ indicates that i, j are to be summed over the first n positive integers excluding τ. The number of solutions of (8) will coincide with $N(p)$. The determinant of the quadratic coefficients of (8) is identically $-b_\tau^{2(n-2)}B$, as may be seen for example by considering this determinant algebraically, subtracting b_i/b_1 times the first column from the column corresponding to the variable x_i for each $i \neq 1$ or τ, then multiplying every column except the first by b_1 and finally dividing every row except the first by $b_1 b_\tau^2$. The determinant of the whole left side of (8) considered as a form in variables x, s, which is the determinant corresponding to δ' of lemma, is easily seen to be identically $b_\tau^{2(n-2)}(b_\tau^2 A + a_\tau B)$. The formulas (7) are then obtained directly from the lemma, recalling only that $A \equiv a$, $B \equiv b \pmod p$.

4. **$N(p)$ for unrestricted invariants.** We shall write $\chi_i(x)$, $(i = 1, \cdots, n)$, for the ith concomitant* of a quadratic form $\chi(x) = \chi_1(x)$ and $(\chi_i \mid p)$ for the quadratic character (mod p) of any integer prime to p represented by $\chi_i(x)$, implying that the character is definitely 1 or -1. If $p \mid r$, we also define $\psi(x, t) = f(x) + 2tg(x) + ct^2$ where c is an integer satisfying $rc \equiv s^2 \pmod p$. We can then prove

THEOREM 2. *The number $N(p)$ of solutions of (1^1) is given in all cases by the following table:*

	Hypothesis	Value of $N(p) - p^{n-1}$ (μ even)	Value of $N(p) - p^{n-2}$ (μ odd)
(9_1)	$\nu = \mu + 2$	$p^{n-2-(\mu)}p[r]((-1)^{(\mu)}a(\mu) \mid p)$	$p^{n-2-(\mu)}((-1)^{(\mu)}ra(\mu) \mid p)$
(9_2)	$\nu = \mu + 1$	$p^{n-1-(\mu)}((-1)^{(\mu)}k(\mu) \mid p)$	$p^{n-2-(\mu)}p[k(\mu)]((-1)^{(\mu)+1}b(\mu) \mid p)$
(9_3)	$\nu = \mu$, $p \mid s$	0	0
(9_4)	$\nu = \mu$, $p \nmid s$	$p^{n-1-(\mu)}p[r]((-1)^{(\mu)}a(\mu) \mid p)$	$p^{n-1-(\mu)}((-1)^{(\mu)}ra(\mu) \mid p)$

Here (μ) in indices is written for $\mu/2$ or $(\mu-1)/2$, according as μ is even or odd; $p[z]$ is written for $p - 1 - p(z^2 \mid p)$ and hence for $p-1$ or -1, according as $p \mid z$ or $p \nmid z$; $a(\mu)$ is defined below, but $(a(\mu) \mid p) = (f_\mu \mid p)$; $b(\mu)$ is defined below, but $(b(\mu) \mid p) = (\phi_{\mu+1} \mid p)$, if $\nu = \mu + 1$; $k(\mu)$ is defined below, but $(k(\mu) \mid p)$ has the following invariant meaning if $\nu = \mu + 1$: if $p \mid r$, $(k(\mu) \mid p) = (s^2 a(\mu) \mid p)$, if $p \nmid r$, $(k(\mu) \mid p) = 0$ or $(r\psi_{\mu+1} \mid p)$, according as $p \mid \psi_{\mu+1}(x)$ or $p \nmid \psi_{\mu+1}(x)$. The other symbols are defined in §1.

* Cf. H. J. S. Smith, *Collected Mathematical Papers*, vol. 1, p. 412.

We may replace (1^1) by (2^1) and define $a(\mu)=a_1a_2\cdots a_\mu$, $b(\mu)=-(b_1^2a_2a_3\cdots a_\mu+b_2^2a_3a_4\cdots a_\mu a_1+\cdots+b_\mu^2a_1a_2\cdots a_{\mu-1})$, $k(\mu)=s^2a(\mu)+rb(\mu)$. Recalling that $p\nmid a(\mu)$, $p\mid a_{\mu+1}, a_{\mu+2}, \cdots, a_n$, a moment's consideration of the matrix of Φ shows that $\mu\leqq\nu\leqq\mu+2$ and that the condition $\nu=\mu+2$ is necessary and sufficient for there to be an index $\tau>\mu$ such that $p\nmid b_\tau$. Hence, if $\nu=\mu+2$, $N(p)$ is $p^{n-\mu-1}$ times the number of solutions of

$$a_1x_1^2+a_2x_2^2+\cdots+a_\mu x_\mu^2\equiv r \pmod p,$$

and taking this number from the note under §2, we have (9_1) valid for $\mu\geqq 1$; if $\nu<\mu+2$, $N(p)$ is $p^{n-\mu}$ times the number of solutions of

$$\begin{aligned} a_1x_1^2+a_2x_2^2+\cdots+a_\mu x_\mu^2&\equiv r,\\ b_1x_1+b_2x_2+\cdots+b_\mu x_\mu&\equiv s \pmod p,\end{aligned}\tag{10}$$

and we can take this number directly from Theorem 1 since the subcases $\nu=\mu+1$, $\nu=\mu$ coincide with the conditions $p\nmid b(\mu)$, $p\mid b(\mu)$, respectively. This gives (9_2), (9_3) and (9_4) for $\mu\geqq 2$. If $\mu=1$ with $\nu=\mu$, every b_i is divisible by p. To prove (9_2) for $\mu=1$, note that $N(p)$ is p^{n-1} times the number of solutions x_1 of $a_1x_1^2\equiv r$, $b_1x_1\equiv s \pmod p$ where $p\nmid a_1b_1$; hence, $N(p)=p^{n-1}\{1-(v^2\mid p)\}$ where $v=a_1s^2-b_1^2r$, a formula to which (9_2) reduces for $\mu=1$.

The statements regarding the invariant values of $(a(\mu)\mid p)$, $(b(\mu)\mid p)$ and $(k(\mu)\mid p)$ may be justified as follows. By definition, $a(\mu)$ is the only nonzero (mod p) determinant of order μ in the matrix of F; $b(\mu)$, in the case $\nu=\mu+1$, is the only nonzero (mod p) determinant of order $\mu+1$ in the matrix of Φ; $k(\mu)/r$, in the case $\nu=\mu+1$, $p\nmid r$, is a determinant of order $\mu+1$ in the matrix of $\Psi(x,t)=F(x)+2tG(x)+ct^2$ while all the other determinants of like order certainly vanish (mod p). Also each of these three determinants is principal. These statements remain true if f' is substituted for F in each of the three forms; but the three forms resulting are equivalent respectively to f, ϕ and ψ. Recalling then that corresponding concomitants of two equivalent forms are likewise equivalent, the statements of the theorem are seen to be correct.

5. **Modulus p^m.** With $m>1$ it seems we have to distinguish between ordinary and singular solutions (cf. §1).

THEOREM 3. *With $m\geqq 1$, the number $M(p^m)$ of ordinary solutions of (1^m) is $p^{(n-2)(m-1)}M(p)$, where $M(p)$ is the number of ordinary solutions of (1^1).*

We may replace (1^m) by (2^m). The theorem being clearly true for

$m=1$, let $m>1$. Every ordinary solution of (2^m) is an ordinary solution of (2^{m-1}) and hence each of the former is represented just once by $z_i=x_i+p^{m-1}y_i$, $(i=1, 2, \cdots, n)$, as x ranges through a complete set of ordinary solutions of (2^{m-1}) and, with each such x, y ranges through a complete (mod p) set of solutions of the congruences for y obtained by substituting z for the variables in (2^m). These congruences are

$$(11) \qquad \sum_1^n a_i x_i y_i \equiv \rho, \qquad \sum_1^n b_i y_i \equiv \sigma \pmod{p},$$

where ρ, σ are integers dependent on x. The condition that x be not singular implies that the matrix of the coefficients of y in (11) has modulo p rank equal to 2; hence, (11) has precisely p^{n-2} solutions y. The theorem follows immediately.

In cases where singular solutions occur, formulas for the total number $N(p^m)$ of solutions of (1^m) are not yet available except for $m=1$. The following theorem, taken with the two preceding, gives explicit formulae for $M(p^m)$ in all cases and for $N(p^m)$ in the cases where no singular solutions occur.

THEOREM 4. *The number of singular solutions of* (1^1) *is* $p^{n+1-\nu}$ *in each of the three cases*: (i) $\nu=\mu+2$, $p\,|\,r$; (ii) $\nu=\mu+1$, $(k(\mu)\,|\,p)=0$; (iii) $\nu=\mu$, $p\,|\,r$, $p\,|\,s$. *In no other cases have the congruences singular solutions.* (Cf. Theorem 2 for definition of $k(\mu)$ and invariant interpretation of $(k(\mu)\,|\,p)$.)

We consider congruences (2^1). The existence of a singular solution x implies that of an integer λ satisfying

$$(12^n) \qquad a_i x_i \equiv \lambda b_i, \qquad i = 1, 2, \cdots, n,$$

$$(13) \qquad r \equiv \lambda s \pmod{p},$$

where (13) is obtained by substituting λb_i for each $a_i x_i$ in (2^1) which gives $\lambda\sum b_i x_i \equiv r$, $\sum b_i x_i \equiv s \pmod{p}$.

Now let $\nu=\mu+2$. There is a subscript $\tau>\mu$ such that $p\nmid b_\tau$, $p\,|\,a_\tau$. Formulas (12) and (13) then imply that $\lambda\equiv r\equiv 0 \pmod{p}$. Conversely, if $r\equiv 0$, choosing $\lambda\equiv 0$ determines $x_1\equiv x_2\equiv \cdots \equiv x_\mu\equiv 0$ and leaves x_i arbitrary for $i>\mu$; to be a singular solution of (2^1), this set must also satisfy $\sum_{\mu+1}^n b_i x_i\equiv s \pmod{p}$ and these conditions suffice. Thus there are $p^{n-\mu-1}$ singular solutions.

Let $\nu<\mu+2$. There is no such subscript τ and (2^1) admits precisely $p^{n-\mu}$ times as many singular solutions as (10). But if $(x_1, x_2, \cdots, x_\mu)$ is a singular solution of (10), (12^μ) and (13) must be satisfied and

substitution from (12^{μ}) in the linear congruence of (10) yields $\lambda b(\mu)+sa(\mu)\equiv 0$; this, multiplied by s, gives $p \mid k(\mu)$ in view of (13). If $\nu=\mu$ (that is, if $p \mid b(\mu)$), this further implies $p \mid s, p \mid r$. Conversely let $p \mid k(\mu)$. If $p \nmid s$, the singular solution of (10) is uniquely determined by (12^{μ}) and (13); if $p \mid s$, then by (13) $p \mid r$. Taking $(x_1, x_2, \cdots, x_{\mu})$ from (12^{μ}), the left members of (10) are congruent respectively to $\lambda^2 b(\mu)/a(\mu)$ and $\lambda b(\mu)/a(\mu)$. Thus (10) is satisfied by singular solutions in p ways (λ arbitrary) if $p \mid b(\mu)$ and uniquely ($\lambda \equiv 0$) if $p \nmid b(\mu)$.

6. **Simple applications.** It follows simply from this theorem that, *if* $p \nmid k = s^2 a + rb$, *then* (1^1) *admits no singular solution.* (For if $p \nmid a$, then $\mu=n$, $\nu \leqq \mu+1$, and $k=k(\mu)$; while if $p \mid a$ the condition implies $p \nmid rb$ and hence $\nu=\mu+2$, $p \nmid r$.) Thus Theorems 1 and 3 give very simple formulas for $N(p^m)$ in this case. Other similar conclusions can easily be drawn.

Theorem 1 has the following interesting application, pointed out by Dr. Gordon Pall. *Let* a_{ij} *be a symmetric matrix of integers, of order 3 and with determinant prime to the odd, square-free integer* m. *Then a necessary and sufficient condition that two solutions* x, y *of the congruence*

$$\sum a_{ij}\xi_i\xi_j \equiv 0 \pmod{m} \tag{14}$$

should satisfy

$$\sum a_{ij}x_iy_j \equiv 0 \pmod{m} \tag{15}$$

is that x, y *be linearly dependent* (mod m).

For, taking m equal to a prime p, the number of simultaneous solutions y of the quadratic congruence (14), with y in place of ξ, and the linear congruence (15), for $x \not\equiv 0$, is easily calculated by the formula (7_2) to be p; and this number is exhausted by the solutions $y \equiv (\lambda x_1, \lambda x_2, \lambda x_3) \pmod{p}$, $\lambda = 1, 2, \cdots, p$. This is easily extended to m as specified by the Chinese Remainder Theorem.

If we relate integral vectors to the quadratic form with matrix (a_{ij}) and adjoint (A_{ij}) by defining norm of $x=\sum A_{ij}x_ix_j$, inner product of x, $y=\sum A_{ij}x_iy_j$, it follows (since $|A_{ij}|$ with $|a_{ij}|$ is prime to m) that *a necessary and sufficient condition that two vectors of norm zero* (mod m) *be linearly dependent* (mod m) *is that their inner product be zero* (mod m).

WESTON COLLEGE

A NOTE ON REICHENBACH'S AXIOMS FOR PROBABILITY IMPLICATION

J. C. C. McKINSEY

In Hans Reichenbach's book *Wahrscheinlichkeitslehre*, the following axioms,* among others, are asserted for the relation of "probability implication":

I. $(p \neq q) \supset [(O \rightarrow_p P) \cdot (O \rightarrow_q P) \equiv (\bar{O})]$.†

II2. $(O \rightarrow_p P) \supset (p \geqq 0)$.

III. $(O \rightarrow_p P) \cdot (O \rightarrow_q Q) \cdot (O \cdot P \supset \bar{Q}) \supset (O \rightarrow_r P \vee Q) \cdot (r = p + q)$.

The proposal is made by Reichenbach, that these axioms be added to a system of logic. The exact character of this system of logic is not specified, but we are presumably to suppose that it is something like the system of *Principia Mathematica*. I must refer the reader to Reichenbach's book for an explanation of the notation occurring in these axioms. Reichenbach does not explicitly state the range of variation of the variables p, q, r, and so on; I shall suppose he intends that these variables can assume as values any real numbers,‡ including also negative real numbers, and positive real numbers greater than $+1$.

I shall now show that these axioms lead to a contradiction.

From Axiom I, we can easily derive the following:

(1) $(\bar{O}) \supset (O \rightarrow_p P)$.

(This is stated as a theorem by Reichenbach on p. 67.) From (1) we get, by substitution,

(2) $(\overline{O \cdot O}) \supset (O \cdot \bar{O} \rightarrow_p P)$.

Reichenbach defines (on p. 67), the expression $(\bar{O})$ as follows:

(3) $(\bar{O}) = (i)(\overline{x_i \in O})$.

* See p. 65 and p. 69.

† For typographical reasons, I express the proposition "O implies P with probability of degree p" by the symbolism "$O \rightarrow_p P$," instead of by the symbolism of Reichenbach.

‡ We might, on the other hand, suppose that these variables can assume as values only real numbers from the closed interval (0, 1). It is not very plausible, however, to suppose this is what Reichenbach intends; such a supposition, moreover, leads in turn to difficulties.

From (3) we have, by substitution,

$$(4) \qquad \overline{(O \cdot \bar{O})} = (i)(\overline{x_i \in O \cdot \bar{O}}).$$

Since the right member of (4) is a provable sentence, the left member is provable also, and we have

$$(5) \qquad \overline{(O \cdot \bar{O})}.$$

From (5) and (2), by modus ponens, we derive

$$(6) \qquad (O \cdot \bar{O} \rightarrow_p P).$$

If now we replace O by $O \cdot \bar{O}$ in Axiom II2, we have

$$(7) \qquad (O \cdot \bar{O} \rightarrow_p P) \supset (p \geqq 0).$$

From (6) and (7), by modus ponens, we have

$$(8) \qquad p \geqq 0.$$

Since it is being assumed that the variable p can assume as a value any real number (and thus, in particular, a negative real number), it is seen that (8) is a contradiction.

This contradiction can be avoided by modifying Axiom I as follows:

$$\text{I}'. \qquad [(p \geqq 0) \cdot (1 \geqq p) \cdot (q \geqq 0) \cdot (1 \geqq q) \cdot (p \neq q)] \supset [(O \rightarrow_p P) \cdot (O \rightarrow_q P) \equiv (\bar{O})].$$

Since similar difficulties arise in connection with Axiom III, it would be desirable also to modify this axiom as follows:

$$\text{III}'. \qquad [(p \geqq 0) \cdot (q \geqq 0) \cdot (1 \geqq p + q) \cdot (O \rightarrow_p P) \cdot (O \rightarrow_q Q) \cdot (O \cdot P \supset \bar{Q})] \supset [(O \rightarrow_r P \vee Q) \cdot (r = p + q)].$$

These changes in the axioms would, of course, entail modifying certain of Reichenbach's theorems. It would be found necessary, for example, to add as an additional hypothesis to certain theorems, conditions like $1 \geqq p$, or $q \geqq 0$, or $1 \geqq p+q$, and so on. It would also be desirable* to add to some of Reichenbach's theorems the hypothesis $\overline{(O)}$.

NEW YORK UNIVERSITY

* This fact was called to my attention by the referee.